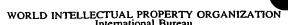
BEST AVAILABLE COPY









INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6:

C12N 1/21, 5/10, 9/10, 15/53, 15/54, 15/61, 15/63, C12P 23/00, C12Q 1/68

(11) International Publication Number:

WO 99/63055

A1

(43) International Publication Date:

9 December 1999 (09.12.99)

(21) International Application Number:

PCT/US99/12121

(22) International Filing Date:

2 June 1999 (02.06.99)

(30) Priority Data:

09/088,724

2 June 1998 (02.06.98)

us

09/088,725

2 June 1998 (02.06.98)

US

(71) Applicant (for all designated States except US): UNIVERSITY OF MARYLAND [US/US]; Office of Technology Liaison, 4312 Knox Road, College Park, MD 20742 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): CUNNINGHAM, Francis, X., Jr. [US/US]; 2727 Washington Avenue, Chevy Chase, MD 20815 (US). SUN, Zairen [US/US]; 3405 Tulane Drive #22, Hyattsville, MD 20783 (US).

(74) Agents: GOLDHUSH, Douglas, H. et al.; Nikaido, Marmelstein, Murray & Oram LLP, Suite 330 - G Street Lobby, Metropolitan Square, 655 Fifteenth Street, N.W., Washington, DC 20005-5701 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report. With amended claims.

(54) Title: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND METHODS OF USE THEREOF

(57) Abstract

Nucleic acid sequences encoding ϵ -cyclase, isopentenyl pyrophosphate isomerase and β -carotene hydroxylase as well as vectors containing the same and hosts transformed with the vectors. Methods for controlling the ratio of various carotenoids in a host and for the production of novel carotenoid pigments. The present invention also provides a method for screening for eukaryotic genes encoding carotenoid biosynthesis, and for modifying the disclosed enzymes.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| AL | Albania | ES | Spain | LS | Lesotho | SI | Slovenia |
|----|--------------------------|------|---------------------|----|-----------------------|-----|-------------------------|
| AM | Armenia | FI | Finland | LT | Lithuania | SK | Slovakia |
| ΑT | Austria | FR | France | LU | Luxembourg | SN | Senegal |
| ΑU | Australia | GA | Gabon | LV | Latvia | SZ | Swaziland |
| AZ | Azerbaijan | GB | United Kingdom | MC | Monaco | TD | Chad |
| BA | Bosnia and Herzegovina | GE | Georgia | MD | Republic of Moldova | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagascar | TJ | Tajikistan |
| BE | Belgium | GN | Guinea | MK | The former Yugoslav | TM | Turkmenistan |
| BF | Burkina Faso | GR | Greece | | Republic of Macedonia | TR | Turkey |
| BG | Bulgaria | HU | Hungary | ML | Mali | TT | Trinidad and Tobago |
| BJ | Benin | IE | Ireland | MN | Mongolia | UA | Ukraine |
| BR | Brazil | IIL. | Israel | MR | Mauritania | UG | Uganda |
| BY | Belarus | IS | Iceland | MW | Malawi | US | United States of Americ |
| CA | Canada | IT | Italy | MX | Mexico | UZ | Uzbekistan |
| CF | Central African Republic | JP | Japan | NE | Niger | VN | Viet Nam |
| CG | Congo | KE | Kenya | NL | Netherlands | YU | Yugoslavia |
| CH | Switzerland | KG | Kyrgyzstan | NO | Norway | zw | Zimbabwe |
| CI | Côte d'Ivoire | KP | Democratic People's | NZ | New Zealand | 2., | 2020 W C |
| CM | Cameroon | | Republic of Korea | PL | Poland | | |
| CN | China | KR | Republic of Korea | PT | Portugal | | |
| CU | Cuba | KZ | Kazakstan | RO | Romania | | |
| CZ | Czech Republic | LC | Saint Lucia | RU | Russian Federation | | |
| DE | Germany | LI | Liechtenstein | SD | Sudan | | |
| DK | Denmark | LK | Sri Lanka | SE | Sweden | | |
| EE | Estonia | LR | Liberia | SG | Singapore | | |



GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND METHODS OF USE THEREOF

BACKGROUND OF THE INVENTION

Field of the Invention

5

10

The present invention describes nucleic acid sequences for eukaryotic genes encoding ϵ lycopene ϵ -cyclase (also known as ϵ -cyclase and ϵ lycopene cyclase), isopentenyl pyrophosphate isomerase (IPP) and β -carotene hydroxylase as well as vectors containing the same and hosts transformed with said vectors. The present invention also provides methods for augmenting the accumulation of carotenoids, changing the composition of the carotenoids, and producing novel and rare carotenoids. The present invention provides methods for controlling the ratio or relative amounts of various carotenoids in a host. The invention also relates to modified lycopene ϵ -cyclase, IPP isomerase and β -carotene hydroxylase. Additionally, the present invention provides a method for screening for genes and cDNAs encoding enzymes of carotenoid biosynthesis and metabolism.

15

20

25

Background of the Invention

Carotenoid pigments with cyclic endgroups are essential components of the photosynthetic apparatus in oxygenic photosynthetic organisms (e.g., cyanobacteria, algae and plants; Goodwin, 1980). The symmetrical bicyclic yellow carotenoid pigment βcarotene (or, in rare cases, the asymmetrical bicyclic \alpha-carotene) is intimately associated with the photosynthetic reaction centers and plays a vital role in protecting against potentially lethal photooxidative damage (Koyama, 1991). β-carotene and other carotenoids derived from it or from α -carotene also serve as light-harvesting pigments (Siefermann-Harms, 1987), are involved in the thermal dissipation of excess light energy captured by the lightharvesting antenna (Demmig-Adams & Adams, 1992), provide substrate for the biosynthesis of the plant growth regulator abscisic acid (Rock & Zeevaart, 1991; Parry & Horgan, 1991), and are precursors of vitamin A in human and animal diets (Krinsky, 1987). Plants also exploit carotenoids as coloring agents in flowers and fruits to attract pollinators and agents of seed dispersal (Goodwin, 1980). The color provided by carotenoids is also of agronomic value in a number of important crops. Carotenoids are currently harvested from a variety of organisms, including plants, algae, yeasts, cyanobacteria and bacteria, for use as pigments in food and feed.

30

The probable pathway for formation of cyclic carotenoids in plants, algae and cyanobacteria is illustrated in Figure 1. Two types of cyclic endgroups or rings are commonly found in higher plant carotenoids, these are referred to as the β (beta) and ϵ (epsilon) rings (Fig. 3). The precursor acyclic endgroup (no ring structure) is referred to as the Ψ (psi) endgroup. The β and ϵ endgroups differ only in the position of the double bond in the ring. Carotenoids with two β rings are ubiquitous, and those with one β and one ϵ ring are common, but carotenoids with two ϵ rings are uncommon. β -carotene (Fig. 1) has two β -endgroups and is a symmetrical compound that is the precursor of a number of other important plant carotenoids such as zeaxanthin and violaxanthin (Fig. 2).

10

15

20

25

5

Genes encoding enzymes of carotenoid biosynthesis have previously been isolated from a variety of sources including bacteria (Armstrong et al., 1989, Mol. Gen. Genet. 216, 254-268; Misawa et al., 1990, J. Bacteriol., 172, 6704-12), fungi (Schmidhauser et al., 1990, Mol. Cell. Biol. 10, 5064-70), cyanobacteria (Chamovitz et al., 1990, Z. Naturforsch, 45c, 482-86; Cunningham et al., 1994) and higher plants (Bartley et al., Proc. Natl. Acad. Sci USA 88, 6532-36; Martinez-Ferez & Vioque, 1992, Plant Mol. Biol. 18, 981-83). Many of the isolated enzymes show a great diversity in structure, function and inhibitory properties between sources. For example, phytoene desaturases from the cyanobacterium Synechococcus and from higher plants and green algae carry out a two-step desaturation to yield ζ-carotene as a reaction product. In plants and cyanobacteria a second enzyme (ζcarotene desaturase), similar in amino acid sequence to the phytoene desaturase, catalyzes two additional desaturations to yield lycopene. In contrast, a single desaturase enzyme from Erwinia herbicola and from other bacteria introduces all four double bonds required to form lycopene. The Erwinia and other bacterial desaturases bear little amino acid sequence similarity to the plant and cyanobacterial desaturase enzymes, and are thought to be of unrelated ancestry. Therefore, even with a gene in hand from one source, it may be difficult to identify a gene encoding an enzyme of similar function in another organism. In particular, the sequence similarity between certain of the prokaryotic and eukaryotic genes encoding enzymes of carotenoid biosynthesis is quite low.

30

Further, the mechanism of gene expression in prokaryotes and eukaryotes appears to differ sufficiently such that one cannot expect that an isolated eukaryotic gene will be properly expressed in a prokaryotic host.



The difficulties in isolating genes encoding enzymes with similar functions is exemplified by recent efforts to isolate the gene encoding the enzyme that catalyzes the formation of β -carotene from the acyclic precursor lycopene. Although a gene encoding an enzyme with this function had been isolated from a bacterium, it had not been isolated from any photosynthetic procaryote or from any eukaryotic organism. The isolation and characterization of the enzyme catalyzing formation of β -carotene in the cyanobacterium Synechococcus PCC7942 was described by the present inventors and others (Cunningham et al., 1993 and 1994). The amino acid sequence similarity of the cyanobacterial enzyme to the various bacterial lycopene β -cyclases is so low (ca. 18-25% overall; Cunningham et al., 1994) that there is much uncertainty as to whether they share a common ancestry or, instead, represent an example of convergent evolution.

The need remains for the isolation of eukaryotic and prokaryotic genes and cDNAs encoding polypeptides involved in the carotenoid biosynthetic pathway, including those encoding a lycopene ϵ -cyclase, IPP isomerase and β -carotene hydroxylase. There remains a need for methods to enhance the production of carotenoids, to alter the composition of carotenoids, and to reduce or eliminate carotenoid production. There also remains a need in the art for methods for screening for genes and cDNAs encoding enzymes of carotenoid biosynthesis and metabolism.

SUMMARY OF THE INVENTION

20

25

30

5

10

15

Accordingly, a first object of this invention is to provide purified and/or isolated nucleic acids which encode enzymes involved in carotenoid biosynthesis; in particular, lycopene ϵ -cyclase, IPP isomerase and β -carotene hydroxylase.

A second object of this invention is to provide purified and/or isolated nucleic acids which encode enzymes which produce novel or uncommon carotenoids.

A third object of the present invention is to provide vectors containing said genes.

A fourth object of the present invention is to provide hosts transformed with said vectors.

Another object of the present invention is to provide hosts which accumulate novel or uncommon carotenoids or which accumulate greater amounts of specific or total carotenoids.

Another object of the present invention is to provide hosts with inhibited and/or altered carotenoid production.

10

15

20

25

30

Another object of this invention is to secure the expression of eukaryotic carotenoidrelated genes in a recombinant prokaryotic host.

Yet another object of the present invention is to provide a method for screening for eukaryotic and prokaryotic genes and cDNAs which encode enzymes involved in carotenoid biosynthesis and metabolism.

An additional object of the invention is to provide a method for manipulating carotenoid biosynthesis in photosynthetic organisms by inhibiting the synthesis of certain enzymatic products to cause accumulation of precursor compounds.

Another object of the invention is to provide modified lycopene ϵ -cyclase, IPP isomerase and β -carotene hydroxylase.

These and other objects of the present invention have been realized by the present inventors as described below.

A subject of the present invention is an isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity and having the amino acid sequence of SEQ ID NOS: 2, 4, 14-21 or 23-27.

The invention also includes vectors which comprise any of the nucleic acid sequences listed above, and host cells transformed with such vectors.

Another subject of the present invention is a method of producing or enhancing the production of a carotenoid in a host cell, comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence to produce the protein.

Yet another subject of the present invention is a method of modifying the production of carotenoids in a host cell, the method comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which produces an RNA and/or encodes for a protein which modifies lycopene ε-cyclase, IPP isomerase or β-carotene hydroxylase enzyme activity, relative to an untransformed host cell, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence in the host cell to modify the production of the carotenoids in the host cell, relative to the untransformed host cell.

10

15

20

25

30



The present invention also includes a method of expressing, in a host cell, a heterologous nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity, the method comprising inserting into the host cell a vector comprising the heterologous nucleic acid sequence, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence.

Also included is a method of expressing, in a host cell, a heterologous nucleic acid sequence which encodes for a protein which modifies lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity in the host cell, relative to an untransformed host cell, the method comprising inserting into the host cell a vector comprising the heterologous nucleic acid sequence, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence.

Another subject of the present invention is a method for screening for genes and cDNAs which encode enzymes involved in carotenoid biosynthesis and metabolism.

BRIEF DESCRIPTION OF THE DRAWINGS

A more complete appreciation of the invention and many of the attendant advantages thereof will be readily obtained as the same becomes better understood by reference to the following detailed description when considered in connection with the accompanying drawings, wherein:

Figure 1 is a schematic representation of the putative pathway of β -carotene biosynthesis in cyanobacteria, algae and plants. The enzymes catalyzing various steps are indicated at the left. Target sites of the bleaching herbicides NFZ and MPTA are also indicated at the left. Abbreviations: DMAPP, dimethylallyl pyrophosphate; FPP, farnesyl pyrophosphate; GGPP, geranylgeranyl pyrophosphate; GPP, geranyl pyrophosphate; IPP, isopentenyl pyrophosphate; LCY, lycopene cyclase; MVA, mevalonic acid; MPTA, 2-(4-methylphenoxy)triethylamine hydrochloride; NFZ, norflurazon; PDS, phytoene desaturase; PSY, phytoene synthase; ZDS, ζ -carotene desaturase; PPPP, prephytoene pyrophosphate.

Figure 2 depicts possible routes of synthesis of cyclic carotenoids and common plant and algal xanthophylls (oxycarotenolds) from neurosporene. Demonstrated activities of the β - and ϵ -cyclase enzymes of A. thaliana are indicated by bold arrows labelled with β or ϵ respectively. A bar below the arrow leading to ϵ -carotene indicates that the enzymatic

10

15

20

25

30



activity was examined but no product was detected. The steps marked by an arrow with a dotted line have not been specifically examined. Conventional numbering of the carbon atoms is given for neurosporene and α -carotene. Inverted triangles (∇) mark positions of the double bonds introduced as a consequence of the desaturation reactions.

Figure 3 depicts the carotene endgroups which are found in plants.

Figure 4 is a DNA sequence and the predicted amino acid sequence of a lycopene ϵ -cyclase cDNA isolated from A. thaliana (SEQ ID NOS: 1 and 2). These sequences were deposited under Genbank accession number U50738. This cDNA is incorporated into the plasmid pATeps.

Figure 5 is a DNA sequence encoding the β -carotene hydroxylase isolated from A. thaliana (SEQ ID NO: 3). This cDNA is incorporated into the plasmid pATOHB.

Figure 6 is an alignment of the predicted amino acid sequences of A. thaliana β-carotene hydroxylase (SEQ ID NO: 4) with those of the bacterial β-carotene hydroxylase enzymes from Alicalgenes sp. (SEQ ID NO: 5) (Genbank D58422), Erwinia herbicola Eho10 (SEQ ID NO.: 6) (GenBank M872280), Erwinia uredovora (SEQ ID NO.: 7) (GenBank D90087) and Agrobacterium aurianticum (SEQ ID NO.: 8) (GenBank D58420). A consensus sequence is also shown. All five genes are identical where a capital letter appears in the consensus. A lowercase letter indicates that three of five, including A. thaliana, have the identical residue. TM; transmembrane.

Figure 7 is a DNA sequence of a cDNA encoding an IPP isomerase isolated from A. thaliana (SEQ ID NO: 9). This cDNA is incorporated into the plasmid pATDP5.

Figure 8 is a DNA sequence of a second cDNA encoding another IPP isomerase isolated from A. thaliana (SEQ ID NO: 10). This cDNA is incorporated into the plasmid pATDP7.

Figure 9 is a DNA sequence of a cDNA encoding an IPP isomerase isolated from *Haematococcus pluvialis* (SEQ ID NO: 11). This cDNA is incorporated into the plasmid pHP04.

Figure 10 is a DNA sequence of a second cDNA encoding another IPP isomerase isolated from *Haematococcus pluvialis* (SEQ ID NO: 12). This cDNA is incorporated into the plasmid pHP05.

Figure 11 is an alignment of the amino acid sequences predicted by IPP isomerase cDNAs isolated from A. thaliana (SEQ ID NO.: 16 and 18), H. pluvialis (SEQ ID NOS.: 14

WO 99/63055 PCT/US99/12121

and 15), Clarkia breweri (SEQ ID NO.: 17) (See, Blanc & Pichersky, Plant Physiol. (1995) 108:855; Genbank accession no. X82627) and Saccharomyces cerevisiae (SEQ ID NO.: 19) (Genbank accession no. J05090).

Figure 12 is a DNA sequence of the cDNA encoding an IPP isomerase isolated from *Tagetes erecta* (marigold; SEQ ID NO: 13). This cDNA is incorporated into the plasmid pPMDP1. xxx's denote a region not originally sequenced. Figure 21A shows the complete marigold sequence.

Figure 13 is an alignment of the consensus sequence of four plant β -cyclases (SEQ ID NO.: 20) with the A. thaliana lycopene ϵ -cyclase (SEQ ID NO.: 21). A capital letter in the plant β consensus is used where all four β -cyclase genes predict the same amino acid residue in this position. A small letter indicates that an identical residue was found in three of the four. Dashes indicate that the amino acid residue was not conserved and dots in the sequence denote a gap. A consensus for the aligned sequences is given, in capital letters below the alignment, where the β - and ϵ -cyclases have the same amino acid residue. Arrows indicate some of the conserved amino acids that will be used as junction sites for construction of chimeric cyclases with novel enzymatic activities. Several regions of interest including a sequence signature indicative of a dinucleotide-binding motif and two predicted transmembrane (TM) helical regions are indicated below the alignment and are underlined.

Figure 14 shows the nucleotide (SEQ ID NO:22) and amino acid sequences (SEQ ID NO:23) of the *Adonis palaestina* (pheasant's eye) ϵ -cyclase cDNA #5.

Figure 15A shows the nucleotide (SEQ ID NO:24) and amino acid sequences (SEQ ID NO:25) of a potato ϵ -cyclase cDNA. Figure 15B shows the amino acid sequence (SEQ ID NO:26) of a chimeric lettuce/potato lycopene ϵ -cyclase. Amino acids in lower case are from the lettuce cDNA and those in upper case are from the potato cDNA. The product of this chimeric cDNA has e-cyclase activity and converts lycopene to the monocyclic δ -carotene.

Figure 16 shows a comparison between the amino acid sequences of the *Arabidopsis* ϵ -cyclase (SEQ ID NO:27) and the potato ϵ -cyclase (SEQ ID NO:25).

Figure 17A shows the nucleotide sequence of the *Adonis palaestina* Ipi1 (SEQ ID NO:28) and Figure 17B shows the nucleotide sequence of the *Adonis palaestina* Ipi2 (SEQ ID NO: 29).

5

10

15

20

25

30

10

15

20

25

30

Figure 18A shows the nucleotide sequence of the *Haematoccus pluvialis* Ipi1 (SEQ ID NO:11) and Figure 18B shows the nucleotide sequence of the *Haematoccus pluvialis* Ipi2 (SEQ ID NO:30).

Figure 19A shows the nucleotide sequence of the *Lactuca sativa (romaine lettuce)* Ipi1 (SEQ ID NO:31) and Figure 19B shows the nucleotide sequence of the *Lactuca sativa* Ipi2 (SEQ ID NO: 32).

Figure 20 shows the nucleotide sequence of the *Chlamydomonas reinhardtii* Ipi1 (SEQ ID NO:33).

Figure 21A shows the nucleotide sequence of the *Tagetes erecta* (marigold) Ipi1 (SEQ ID NO:34) and Figure 21B shows the nucleotide sequence of the *Oryza sativa* (rice) Ipi1 (SEQ ID NO:35).

Figure 22 shows a amino acid sequence alignment of various plant and green algal isopentenyl isomerases (IPI) (SEQ ID NOS:16, 36-45).

Figure 23 shows a comparison between *Adonis palaestina* ϵ -cyclase cDNA #3 and cDNA #5 nucleotide sequences.

Figure 24 shows a comparison between *Adonis palaestina* ϵ -cyclase cDNA #3 and cDNA #5 predicted amino acid sequences.

Figure 25 shows a sequence alignment of various plant β - and ϵ -cyclases. Those sequences outlined in grey denote identical sequences among the ϵ -cyclases. Those sequences outlined in black denote identical sequences among both the β - and ϵ -cyclases.

Figure 26 shows a sequence alignment of the plant ϵ -cyclases from Figure 25. Those sequences outlined in black denote identical sequences among the ϵ -cyclases.

Figure 27 is a dendrogram or "tree" illustrating the degree of amino acid sequence similarity for various lycopene β - and ϵ -cyclases.

Figure 28 shows a comparison between Arabidopsis ϵ -cyclase and lettuce ϵ -cyclase predicted amino acid sequences.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention includes an isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ε-cyclase, IPP isomerase or β-carotene hydroxylase enzyme activity and having the amino acid sequence of SEQ ID NOS: 2, 4, 14-21, 23 or 25-27. Nucleic acids encoding lycopene ε-cyclase, β-carotene hydroxylase and IPP

10

15

20

25

30

isomerases have been isolated from several genetically distant sources.

The present inventors have isolated nucleic acids encoding the enzyme IPP isomerase, which catalyzes the reversible conversion of isopentenyl pyrophosphate (IPP) to dimethylallyl pyrophosphate (DMAPP). IPP isomerase cDNAs were isolated from the plants A. thaliana, Tagetes erecta (marigold), Adonis palaestina (pheasant's eye), Lactuca sativa (romaine lettuce) and from the green algae H. pluvialis and Chlamydomonas reinhardtii.

Alignments of the amino acid sequences predicted by some of these cDNAs are shown in Figures 12 and 22. Plasmids containing some of these cDNAs were deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville MD 20852 on March 4, 1996 under ATCC accession numbers 98000 (pHP05 - H. pluvialis); 98001 (pMDP1 - marigold); 98002 (pATDP7 - A. thaliana) and 98004 (pHP04 - H. pluvialis).

The present inventors have also isolated nucleic acids encoding the enzyme β-carotene hydroxylase, which is responsible for hydroxylating the β-endgroup in carotenoids. The nucleic acid of the present invention is shown in SEQ ID NO: 3 and Figure 5. The full length cDNA product hydroxylates both end groups of β-carotene as do products of cDNAs which encode proteins truncated by up to 50 amino acids from the N-terminus. Products of genes which encode proteins truncated between about 60-110 amino acids from the N-terminus preferentially hydroxylate only one ring. A plasmid containing this gene was deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville MD 20852 on March 4, 1996 under ATCC accession number 98003 (pATOHB - A. thaliana).

The present inventors have also isolated nucleic acids encoding the enzyme lycopene ϵ -cyclase, which is responsible for the formation of ϵ -endgroups in carotenoids. The A. thaliane ϵ -cyclase adds an ϵ ring to only one end of the symmetrical lycopene while the related β -cyclase adds a ring at both ends. The A. thaliana cDNA of the present invention is shown in Figure 4 and SEQ ID NO: 1. A plasmid containing this gene was deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville MD 20852 on March 4, 1996 under ATCC accession number 98005 (pATeps - A. thaliana).

In addition, lycopene ϵ -cyclases have been identified in lettuce and in *Adonis* palaestina (cDNA #5) which encode enzymes that convert lycopene to the bicyclic ϵ -carotene (ϵ , ϵ -carotene). An additional cDNA from *Adonis palaestina* (cDNA #3) encodes a lycopene ϵ -cyclase which converts lycopene into δ -carotene (ϵ , ψ -carotene) and differs from the lycopene ϵ -cyclase which forms bicyclic ϵ -carotene (ϵ , ϵ -carotene) by only 5 amino acids.

One or more of these amino acids may be modified by alteration of the nucleotide sequence in the #5 cDNA to obtain an enzyme which forms the bicyclic ϵ, ϵ -carotene. The sequences of the *Adonis palaestina* and *Arabidopsis thaliana* ϵ -cyclases have about 70% nucleotide identity and about 72% amino acid identity.

10

15

20

5

Initial experiments by the inventors with chimeric genes indicated that the part of the ϵ -cyclase which is responsible for adding 2ϵ rings to form ϵ, ϵ -carotene is the carboxy terminal portion of the gene. The lettuce ϵ -cyclase adds two ϵ rings to form ϵ, ϵ -carotene. A DNA encoding a partial potato ϵ -cyclase (missing its amino terminal portion), when combined with an amino terminal region from the lettuce ϵ -cyclase gene, produces a monocyclic δ -carotene (ϵ, ψ -carotene). With the discovery of the differences between the Adonis palaestina clone #3 and clone #5, the specific amino acids responsible for the addition of an extra ϵ ring have been identified (Figure 24). Specifically, amino acid 55 is Thr in clone #3 and Ser in clone #5, amino acid 210 is Asn in clone #3 and Asp in clone #5, amino acid 231 is Asp in clone #3 and Glu in clone #5, amino acid 352 is Ile in clone #3 and Val in clone #5, and amino acid 524 is Lys in clone #3 and Arg in clone #5. It can be appreciated that these changes are quite conservative, as only one change, at amino acid 210, changes the charge of the protein.

Thus, it is clear that the nucleic acids of the invention encoding the enzymes as presently disclosed may be altered to increase a particularly desirable property of the enzyme, to change a property of the enzyme, or to diminish an undesirable property of the enzyme. Such modifications can be by deletion, substitution, or insertion of one or more amino acids, and can be performed by routine enzymatic manipulation of the nucleic acid encoding the enzyme (such as by restriction enzyme digestion, removal of nucleotides by mung bean nuclease or *Bal*31, insertion of nucleotides by Klenow fragment, and by religation of the ends), by site-directed mutagenesis, or may be accidental, such as by low fidelity PCR or those obtained through mutations in hosts that are producers of the enzymes. These techniques as well as other suitable techniques are well known in the art.

30

25

Mutations can be made in the nucleic acids of the invention such that a particular codon is changed to a codon which codes for a different amino acid. Such a mutation is generally made by making the fewest nucleotide changes possible. A substitution mutation of this sort can be made to change an amino acid in the resulting protein in a non-conservative manner (i.e., by changing the codon from an amino acid belonging to a grouping

15

20

25

30

of amino acids having a particular size or characteristic to an amino acid belonging to another grouping) or in a conservative manner (i.e., by changing the codon from an amino acid belonging to a grouping of amino acids having a particular size or characteristic to an amino acid belonging to the same grouping). Such a conservative change generally leads to less change in the structure and function of the resulting protein. A non-conservative change is more likely to alter the structure, activity or function of the resulting protein. The present invention should be considered to include sequences containing conservative changes which do not significantly alter the activity or binding characteristics of the resulting protein.

The following is one example of various groupings of amino acids:

Amino acids with nonpolar R groups: Alanine, Valine, Leucine, Isoleucine, Proline, Phenylalanine, Tryptophan and Methionine.

Amino acids with uncharged polar R groups: Glycine, Serine, Threonine, Cysteine, Tyrosine, Asparagine and Glutamine.

Amino acids with charged polar R groups (negatively charged at Ph 6.0): Aspartic acid and Glutamic acid.

Basic amino acids (positively charged at pH 6.0): Lysine, Arginine and Histidine.

Another grouping may be those amino acids with phenyl groups: Phenylalanine, Tryptophan and Tyrosine.

Another grouping may be according to molecular weight (i.e., size of R groups). Particularly preferred substitutions are:

- Lys for Arg and vice versa such that a positive charge may be maintained;
- Glu for Asp and vice versa such that a negative charge may be maintained;
- Ser for Thr such that a free -OH can be maintained; and
- Gln for Asn such that a free NH₂ can be maintained.

Amino acid substitutions may also be introduced to substitute an amino acid with a particularly preferable property. For example, a Cys may be introduced to provide a potential site for disulfide bridges with another Cys. A His may be introduced as a particularly "catalytic" site (i.e., His can act as an acid or base and is the most common amino acid in biochemical catalysis). Pro may be introduced because of its particularly planar structure, which induces β -turns in the protein's structure.

It is clear that certain modifications of SEQ ID NOS: 2, 4, 14-21, 23 or 25-27 can take place without destroying the activity of the enzyme. It is noted especially that truncated

versions of the nucleic acids of the invention are functional. For example, several amino acids (from 1 to about 120) can be deleted from the N-terminus of the lycopene ε-cyclases of the invention, and a functional protein can still be produced. This fact is made especially clear from Figure 25, which shows a sequence alignment of several plant ε-cyclases. As can be seen from Figure 25, there is an enormous amount of sequence disparity between amino acid sequences 2 to about 50-70 (depending on the particular sequence, since gaps are present). There is less, but also a substantial amount of, sequence dissimilarity between about 50-70 to about 90-120 (depending on the particular sequence). Thereafter, the sequences are fairly conserved, except for small pockets of dissimilarity between about 275-295 to about 285-305 (depending on the particular sequence), and between about 395-415 to about 410-430 (depending on the particular sequence).

The present inventors have found that the amount of the 5' region present in the nucleic acids of the invention can alter the activity of the enzyme. Instead of diminishing activity, truncating the 5' region of the nucleic acids of the invention may result in an enzyme with a different specificity. Thus, the present invention relates to nucleic acids and enzymes encoded thereby which are truncated to within 0-50, preferably 0-25, codons of the 5' initiation codon of their prokaryotic counterparts as determined by alignment maps as discussed below.

For example, when the cDNA encoding A. thaliana β -carotene hydroxylase was truncated, the resulting enzyme catalyzed the formation of β -cryptoxanthin as the major product and zeaxanthin as minor product; in contrast to its normal production of zeaxanthin.

The present invention is intended to include those nucleic acid and amino acid sequences in which substitutions, deletions, additions or other modifications have taken place, as compared to SEQ ID NOS: 2, 4, 14-21, 23 or 25-27, without destroying the activity of the enzyme. Preferably, the substitutions, deletions, additions or other modifications take place at the 5' end, or any other of those positions which already show dissimilarity between any of the presently disclosed amino acid sequences (see also Figure 25) or other amino acid sequences which are known in the art and which encode the same enzyme (i.e., lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase).

In each case, nucleic acid and amino acid sequence similarity and identity is measured using sequence analysis software, for example, the Sequence Analysis, Gap, or BestFit software packages of the Genetics Computer Group (University of Wisconsin Biotechnology

5

10

15

20

25

30

Center, 1710 University Avenue, Madison, Wisconsin 53705), MEGAlign (DNAStar, Inc., 1228 S. Park St., Madison, Wisconsin 53715), or MacVector (Oxford Molecular Group, 2105 S. Bascom Avenue, Suite 200, Campbell, California 95008). Such software uses algorithms to match similar sequences by assigning degrees of identity to various substitutions, deletions, and other modifications, and includes detailed instructions as to useful parameters, etc., such that those of routine skill in the art can easily compare sequence similarities and identities. An example of a useful algorithm in this regard is the algorithm of Needleman and Wunsch, which is used in the Gap program discussed above. This program finds the alignment of two complete sequences that maximizes the number of matches and minimizes the number of gaps. Another useful algorithm is the algorithm of Smith and Waterman, which is used in the BestFit program discussed above. This program creates an optimal alignment of the best segment of similarity between two sequences. Optimal alignments are found by inserting gaps to maximize the number of matches using the local homology algorithm of Smith and Waterman.

15

10

5

Conservative (i.e. similar) substitutions typically include substitutions within the following groups: glycine and alanine; valine, isoleucine and leucine; aspartic acid, glutamic acid, asparagine and glutamine; serine and threonine; lysine and arginine; and phenylalanine and tyrosine. Substitutions may also be made on the basis of conserved hydrophobicity or hydrophilicity (see Kyte and Doolittle, *J. Mol. Biol.* 157: 105-132 (1982)), or on the basis of the ability to assume similar polypeptide secondary structure (see Chou and Fasman, *Adv. Enzymol.* 47: 45-148 (1978)).

20

25

If comparison is made between nucleotide sequences, preferably the length of comparison sequences is at least 50 nucleotides, more preferably at least 60 nucleotides, at least 75 nucleotides or at least 100 nucleotides. It is most preferred if comparison is made between the nucleic acid sequences encoding the enzyme coding regions necessary for enzyme activity. If comparison is made between amino acid sequences, preferably the length of comparison is at least 20 amino acids, more preferably at least 30 amino acids, at least 40 amino acids or at least 50 amino acids. It is most preferred if comparison is made between the amino acid sequences in the enzyme coding regions necessary for enzyme activity.

30

It should be appreciated that also within the scope of the present invention are nucleic acid sequences encoding lycopene ε-cyclases, IPP isomerases and β-carotene hydroxylases

10

15

20

25

30

which code for enzymes having the same amino acid sequence as SEQ ID NOS: 2, 4, 14-21, 23 or 25-27, but which are degenerate to the nucleic acids specifically disclosed herein.

The amino acid residues described herein are preferred to be in the "L" isomeric form. However, residues in the "D" isomeric form can be substituted for any L-amino acid residue, as long as the desired functional property of immunoglobulin-binding is retained by the polypeptide.

In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Sambrook et al, "Molecular Cloning: A Laboratory Manual" (1989); "Current Protocols in Molecular Biology" Volumes I-III [Ausubel, R. M., ed. (1994)]; "Cell Biology: A Laboratory Handbook" Volumes I-III [J. E. Celis, ed. (1994))]; "Current Protocols in Immunology" Volumes I-III [Coligan, J. E., ed. (1994)]; "Oligonucleotide Synthesis" (M.J. Gait ed. 1984); "Nucleic Acid Hybridization" [B.D. Hames & S.J. Higgins eds. (1985)]; "Transcription And Translation" [B.D. Hames & S.J. Higgins, eds. (1984)]; "Animal Cell Culture" [R.I. Freshney, ed. (1986)]; "Immobilized Cells And Enzymes" [IRL Press, (1986)]; B. Perbal, "A Practical Guide To Molecular Cloning" (1984).

The present invention also includes vectors. Suitable vectors according to the present invention comprise a nucleic acid of the invention encoding an enzyme involved in carotenoid biosynthesis or metabolism and a suitable promoter for the host, and can be constructed using techniques well known in the art (for example Sambrook et al., Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; Ausubel et al., Current Protocols in Molecular Biology, Greene Publishing and Wiley Interscience, New York, 1991). Suitable vectors for eukaryotic expression in plants are described in Frey et al., Plant J. (1995) 8(5):693 and Misawa et al, 1994a; incorporated herein by reference. Suitable vectors for prokaryotic expression include pACYC184, pUC119, and pBR322 (available from New England BioLabs, Bevery, MA) and pTrcHis (Invitrogen) and pET28 (Novagen) and derivatives thereof. The vectors of the present invention can additionally contain regulatory elements such as promoters, repressors, selectable markers such as antibiotic resistance genes, etc.

The nucleic acids encoding the carotenoid enzymes as described above, when cloned into a suitable expression vector, can be used to overexpress these enzymes in a plant

10

15

20

25

30



expression system or to inhibit the expression of these enzymes. For example, a vector containing the gene encoding lycopene ϵ -cyclase can be used to increase the amount of α -carotene and carotenoids derived from α -carotene (such as lutein and α -cryptoxanthin) in an organism and thereby alter the nutritional value, pharmacology and visual appearance value of the organism.

Therefore, the present invention includes a method of producing or enhancing the production of a carotenoid in a host cell, relative to an untransformed host cell, the method comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence to produce the protein.

The present invention also includes a method of modifying the production of carotenoids in a host cell, the method comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which produces an RNA and/or encodes for a protein which modifies lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity, relative to an untransformed host cell, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence in the host cell to modify the production of the carotenoids in the host cell, relative to the untransformed host cell.

The term "modifying the production" means that the amount of carotenoids produced in the host cell can be enhanced, reduced, or left the same, as compared to the untransformed host cell. In accordance with one embodiment of the present invention, the make-up of the carotenoids (i.e., the specific carotenoids produced) is changed vis a vis each other, and this change in make-up may result in either a net gain, net loss, or no net change in the total amount of carotenoids produced in the cell. In accordance with another embodiment of the present invention, the production or the biochemical activity of the carotenoids (or the enzymes which catalyze their formation) is enhanced by the insertion of an enzyme-encoding nucleic acid of the invention. In yet another embodiment of the invention, the production or the biochemical activity of the carotenoids (or the enzymes which catalyze their formation) may be reduced or inhibited by a number of different approaches available to those skilled in the art, including but not limited to such methodologies or approaches as anti-sense (e.g.,

10

15

20

25

30

Gray et al (1992) Plant Mol. Biol. 19:69-87), ribozymes (e.g., Wegener et al (1994) Mol. Gen. Genet. 245:465-470), co-suppression (e.g., Fray and Grierson (1993) Plant Mol. Biol. 22:589-602), targeted disruption of the gene (e.g., Schaefer et al. (1997) Plant J. 11:1195-1206), intracellular antibodies (e.g., Rondon and Marasco (1997) Ann. Rev. Microbiol. 51:257-283) or whatever other approaches rely on the knowledge or availability of the nucleic acid or amino acid sequences of the invention and/or portions thereof, to thereby reduce accumulation of carotenoids with ϵ rings and compounds derived from them (for ϵ -cyclase inhibition), or carotenoids with hydroxylated β rings and compounds derived from them (for β -hydroxylase inhibition), or, in the case if IPP isomerase, accumulation of any isoprenoid compound.

Preferably, at least a portion of the nucleic acid sequences used in the methods, vectors and host cells of the invention codes for an enzyme having an amino acid sequence which is at least 85% identical, preferably at least 90%, at least 95% or completely identical to SEQ ID NOS: 2, 4, 14-21, 23 or 25-27. Sequence identity is determined as noted above. Preferably, sequence additions, deletions or other modifications are made as indicated above, so as to not affect the function of the particular enzyme.

In a preferred embodiment, vectors are manufactured which contain a DNA encoding a eukaryotic IPP isomerase upstream of a DNA encoding a second eukaryotic carotenoid enzyme. The inventors have discovered that inclusion of an IPP isomerase gene increases the supply of substrate for the carotenoid pathway; thereby enhancing the production of carotenoid endproducts, as compared to a host cell which is not transformed with such a vector. This is apparent from the much deeper pigmentation in carotenoid-accumulating colonies of *E. coli* which also contain one of the aforementioned IPP isomerase genes when compared to colonies that lack this additional IPP isomerase gene. Similarly, a vector comprising an IPP isomerase gene can be used to enhance production of any secondary metabolite of dimethylallyl pyrophosphate and/or isopentenyl pyrophosphate (such as isoprenoids, steroids, carotenoids, etc.). The term "isoprenoid" is intended to mean any member of the class of naturally occurring compounds whose carbon skeletons are composed, in part or entirely, of isopentyl C₅ units. Preferably, the carbon skeleton is of an essential oil, a fragrance, a rubber, a carotenoid, or a therapeutic compound, such as paclitaxel.

A vector containing the cDNA encoding a lycopene ϵ -cyclase of the invention, preferably the lettuce lycopene ϵ -cyclase or Adonis ϵ -cyclase #5, can be used to increase the

10

15

20

25

30

amount of bicyclic ∈-carotene in an organism and thereby alter the nutritional value, pharmacology and visual appearance value of the organism. In addition, the transformed organism can be used in the formulation of therapeutic agents, for example in the treatment of cancer (see Mayne et al (1996) FASEB J. 10:690-701; Tsushima et al (1995) Biol. Pharm. Bull. 18:227-233).

An antisense strand of a nucleic acid of the invention can be inserted into a vector. For example, the lycopene ϵ -cyclase gene can be inserted into a vector and incorporated into the genomic DNA of a host, thereby inhibiting the synthesis of ϵ , β -carotenoids (lutein and α -carotene) and enhancing the synthesis of β , β -carotenoids (zeaxanthin and β -carotene).

The present invention also relates to novel enzymes which are encoded by the amino acid sequences of the invention, or portions thereof.

The present invention also relates to novel enzymes which can transform known carotenoids into novel or uncommon products. Currently ϵ -carotene (see Figure 2) and γ -carotene are commonly produced only in minor amounts. As described below, an enzyme can be produced which transforms lycopene to γ -carotene and lycopene to ϵ -carotene. With these products in hand, bulk synthesis of other carotenoids derived from them are possible. For example, ϵ -carotene can be hydroxylated to form lactucaxanthin, an isomer of lutein (one ϵ and one β ring) and zeaxanthin (two β rings) where both endgroups are, instead, ϵ rings.

In addition to novel enzymes produced by truncating the 5' region of known enzymes, as discussed above, novel enzymes which can participate in the formation of unusual carotenoids can be formed by replacing portions of one gene with an analogous sequence from a structurally related gene. For example, β -cyclase and ϵ -cyclase are structurally related (see Figure 13). By replacing a portion of β -lycopene cyclase with the analogous portion of ϵ -cyclase, an enzyme which produces γ -carotene will be produced (one β endgroup). Further, by replacing a portion of the lycopene ϵ -cyclase with the analogous portion of β -cyclase, an enzyme which produces ϵ -carotene will be produced (with some exceptions, such as the lettuce ϵ -cyclase, plant ϵ -cyclases normally produce a compound with one ϵ -endgroup, δ -carotene). Similarly, β -hydroxylase could be modified to produce enzymes of novel function by creation of hybrids with ϵ -hydroxylase.

Host systems according to the present invention can comprise any organism that already produces carotenoids or which has been genetically modified to produce carotenoids.

The IPP isomerase genes are more broadly applicable for enhancing production of any product dependent on DMAPP and/or IPP as a precursor.

Organisms which already produce carotenoids include plants, algae, some yeasts, fungi and cyanobacteria and other photosynthetic bacteria. Transformation of these hosts with vectors according to the present invention can be done using standard techniques such as those described in Misawa et al., (1990) supra; Hundle et al., (1993) supra; Hundle et al., (1991) supra; Misawa et al., (1991) supra; Sandmann et al., supra; and Schnurr et al., supra.

Transgenic organisms can be constructed which include the nucleic acid sequences of the present invention (Bird et al, 1991; Bramley et al, 1992; Misawa et al, 1994a; Misawa et al, 1994b; Cunningham et al, 1993). The incorporation of these sequences can allow the controlling of carotenoid biosynthesis, content, or composition in the host cell. These transgenic systems can be constructed to incorporate sequences which allow for the overexpression of the nucleic acids of the present invention. Transgenic systems can also be constructed containing antisense expression of the nucleic acid sequences of the present invention. Such antisense expression would result in the accumulation of the substrates of the substrates of the enzyme encoded by the sense strand.

A method for screening for eukaryotic genes which encode enzymes involved in carotenoid biosynthesis comprises transforming a prokaryotic host with a nucleic acid which may contain a eukaryotic or prokaryotic carotenoid biosynthetic gene; culturing said transformed host to obtain colonies; and screening for colonies exhibiting a different color than colonies of the untransformed host.

Suitable hosts include E. coli, cyanobacteria such as Synechococcus and Synechocystis, alga and plant cells. E. coli are preferred.

In a preferred embodiment, the above "color complementation" screening protocol can be enhanced by using mutants which are either (1) deficient in at least one carotenoid biosynthetic gene or (2) overexpress at least one carotenoid biosynthetic gene. In either case, such mutants will accumulate carotenoid precursors.

Prokaryotic and eukaryotic DNA or cDNA libraries can be screened in total for the presence of genes of carotenoid biosynthesis, metabolism and degradation. Preferred organisms to be screened include photosynthetic organisms.

5

10

15

20

25

30

10

15

20

25

30

E. coli can be transformed with these eukaryotic cDNA libraries using conventional methods such as those described in Sambrook et al, 1989 and according to protocols described by the vendors of the cloning vectors.

For example, the cDNA libraries in bacteriophage vectors such as lambdaZAP (Stratagene) or lambda ZIPLOX (Gibco BRL) can be excised en masse and used to transform *E.coli*.

Transformed *E. coli* can be cultured using conventional techniques. The culture broth preferably contains antibiotics to select and maintain plasmids. Suitable antibiotics include penicillin, ampicillin, chloramphenicol, etc. Culturing is typically conducted at 15-40°C, preferably at room temperature or slightly above (18-28°C), for 12 hours to 7 days.

Cultures are plated and the plates are screened visually for colonies with a different color than the colonies of the host $E.\ coli$ transformed with the empty plasmid cloning vector. For example, $E.\ coli$ transformed with the plasmid, pAC-BETA (described below), produce yellow colonies that accumulate β -carotene. After transformation with a cDNA library, colonies which contain a different hue than those formed by $E.\ coli/pAC$ -BETA would be expected to contain enzymes which modify the structure or accumulation of β -carotene. Similar $E.\ coli$ strains can be engineered which accumulate earlier products in carotenoid biosynthesis, such as lycopene, γ -carotene, etc.

Having generally described this invention, a further understanding can be obtained by reference to certain specific examples which are provided herein for purposes of illustration only and are not intended to be limiting unless otherwise specified.

EXAMPLE

I. <u>Isolation of β-carotene hydroxylase</u>

Plasmid Construction

An 8.6kb BgIII fragment containing the carotenoid biosynthetic genes of *Erwinia herbicola* was first cloned in the BamHI site of plasmid vector pACYC184 (chloramphenicol resistant), and then a 1.1kb BamHI fragment containing the *E. herbicola* β-carotene hydroxylase (*CrtZ*) was deleted. *E.coli* strains containing the resulting plasmid, pAC-BETA, accumulate β-carotene and form yellow colonies (Cunningham et al., 1994).

A full length cDNA encoding IPP isomerase of *Haematococcus pluvialis* (HP04) was first excised with *BamH*I and *Kpn*I from pBluescript SK-, and then ligated into the

corresponding sites of the pTrcHisA vector with high-level expression from the *trc* promoter (Invitrogen, Inc.). A fragment containing the IPP isomerase and *trc* promoter was subsequently excised with *EcoRV* and *KpnI*, treated with the Klenow fragment of DNA polymerase to produce blunt ends, and ligated in the Klenow-treated *HindIII* site of pAC-BETA. *E.coli* cells transformed with this new plasmid pAC-BETA-04 form orange colonies on LB plates (*vs.* yellow for those containing pAC-BETA) and cultures accumulate substantially more β-carotene (*ca.* two fold) than those that contain pAC-BETA.

Screening of an Arabidopsis cDNA Library

Several λ cDNA expression libraries of *Arabidopsis* were obtained from the *Arabidopsis* Biological Resource Center (Ohio State University, Columbus, OH) (Kieber et al., 1993). The λ cDNA libraries were excised *in vivo* using Stratagene's ExAssist SOLR system to produce a phagemid cDNA library wherein each phagemid contained also a gene conferring resistance to the antibiotic ampicillin.

E.coli strain DH10BZIP was chosen as the host cell for the screening and pigment production, although we have also used TOP10F' and XL1-Blue for this purpose. DH10B cells were transformed with plasmid pAC-BETA-04 and were plated on LB agar plates containing chloramphenicol at 50 µg/ml (from United States Biochemical Corporation). The phagemid Arabidopsis cDNA library was then introduced into DH10B cells already containing pAC-BETA-04. Transformed cells containing both pAC-BETA-04 and Arabidopsis cDNA library phagemids were selected on chloramphenicol plus ampicillin (150 µg/ml) agar plates. Maximum color development occurred after 3 to 7 days incubation at room temperature, and the rare bright yellow colonies were selected from a background of many thousands of orange colonies on each agar plate. Selected colonies were inoculated into 3 ml liquid LB medium containing ampicillin and chloramphenicol, and cultures were incubated at room temperature for 1-2 days, with shaking. Cells were then harvested by centrifugation and extracted with acetone in microfuge tubes. After centrifugation, the pigmented extract was spotted onto silica gel thin-layer chromatography (TLC) plates, and developed with a hexane:ether (1:1, by volume) mobile phases. B-carotene hydroxylaseencoding cDNAs were identified based on the appearance of a yellow pigment that comigrated with zeaxanthin on the TLC plates.

5

10

15

20

25

30



Subcloning and Sequencing

The plasmid containing the β-carotene hydroxylase cDNA was recovered and analyzed by standard procedures (Sambrook et al., 1989). The *Arabidopsis* β-carotene hydroxylase was sequenced completely on both strands on an automatic sequencer (Applied Biosystems, Model 373A, Version 2.0.1S). The cDNA insert of 0.95kb also was excised and ligated into the a pTrcHis vector. A *BgI*II restriction site within the cDNA was used to remove that portion of the cDNA that encodes the predicted polypeptide N terminal sequence region that is not also found in bacterial β-carotene hydroxylases (Figure 6). A BgIII-XhoI fragment was directionally cloned in BamHI-XhoI digested TrcHis vectors.

Pigment Analysis

5

10

15

20

25

30

A single colony was used to inoculate 50 ml of LB containing ampicillin and chloramphenicol in a 250-ml flask. Cultures were incubated at 28°C for 36 hours with gentle shaking, and then harvested at 5000 rpm in an SS-34 rotor. The cells were washed once with distilled H₂O and resuspended with 0.5 ml of water. The extraction procedures and HPLC were essentially as described previously (Cunningham et al, 1994).

II. Isolation and biochemical analysis of an Arabidopsis lycopene ϵ -cyclase Plasmid Construction

Construction of plasmids pAC-LYC, pAC-NEUR, and pAC-ZETA is described in Cunningham et al., (1994). In brief, the appropriate carotenoid biosynthetic genes from *Erwinia herbicola*, *Rhodobacter capsulatus*, and *Synechococcus* sp. strain PCC7942 were cloned in the plasmid vector pACYC184 (New England BioLabs, Beverly, MA). Cultures of *E. coli* containing the plasmids pAC-ZETA, pAC-NEUR, and pAC-LYC, accumulate ζ-carotene, neurosporene, and lycopene, respectively. The plasmid pAC-ZETA was constructed as follows: an 8.6-kb BgIII fragment containing the carotenoid biosynthetic genes of *E. herbicola* (GenBank M87280; Hundle et al., 1991) was obtained after partial digestion of plasmid pPL376 (Perry et al., 1986; Tuveson et al., 1986) and cloned in the BamHI site of pACYC184 to give the plasmid pAC-EHER. Deletion of adjacent 0.8- and 1.1-kb BamHI-BamHI fragments (deletion Z in Cunningham et al., 1994), and of a 1.1 kB Sall-Sall fragment (deletion X) served to remove most of the coding regions for the *E. herbicola* β-carotene hydroxylase (crtZ gene) and zeaxanthin glucosyltransferase (crtX gene), respectively. The

10

15

20

25

resulting plasmid, pAC-BETA, retains functional genes for geranylgeranyl pyrophosphate synthase (crtE), phytoene synthase (crtB), phytoene desaturase (crtI), and lycopene cyclase (crtY). Cells of E. coli containing this plasmid form yellow colonies and accumulate β -carotene. A plasmid containing both the lycopene ϵ - and β -cyclase cDNAs of A. thaliana was constructed by excising the ϵ -cyclase in clone y2 as a PvuI-PvuII fragment and ligating this piece in the SnaBI site of a plasmid (pSPORT 1 from GIBCO-BRL) that already contained the β -cyclase (Cunningham et al., 1996).

Organisms and Growth Conditions

E. coli strains TOP10 and TOP10 F' (obtained from Invitrogen Corporation, San Diego, CA) and XL1-Blue (Stratagene) were grown in Luria-Bertani (LB) medium (Sambrook et al., 1989) at 37°C in darkness on a platform shaker at 225 cycles per min. Media components were from Difco (yeast extract and tryptone) or Sigma (NaCl). Ampicillin at 150 μg/mL and/or chloramphenicol at 50 μg/mL (both from United States Biochemical Corporation) were used, as appropriate, for selection and maintenance of plasmids.

Mass Excision and Color Complementation Screening of an A. thaliana cDNA Library

A size-fractionated 1-2 kB cDNA library of A. thaliana in lambda ZAPII (Kieber et al., 1993) was obtained from the Arabidopsis Biological Resource Center at The Ohio State University (stock number CD4-14). Other size fractionated libraries were also obtained (stock numbers CD4-13, CD4-15, and CD4-16). An aliquot of each library was treated to cause a mass excision of the cDNAs and thereby produce a phagemid library according to the instructions provided by the supplier of the cloning vector (Stratagene; E. coli strain XL1-Blue and the helper phage R408 were used). The titre of the excised phagemid was determined and the library was introduced into a lycopene-accumulating strain of E. coli TOP10 F' (this strain contained the plasmid pAC-LYC) by incubation of the phagemid with the E. coli cells for 15 min at 37°C. Cells had been grown overnight at 30°C in LB medium supplemented with 2% (w/v) maltose and 10 mM MgSO₄ (final concentration), and harvested in 1.5 ml microfuge tubes at a setting of 3 on an Eppendorf microfuge (5415C) for 10 min. The pellets were resuspended in 10 mM MgSO₄ to a volume equal to one-half that of the

initial culture volume. Transformants were spread on large (150 mm diameter) LB agar petri plates containing antibiotics to provide for selection of cDNA clones (ampicillin) and maintenance of pAC-LYC (chloramphenicol). Approximately 10,000 colony forming units were spread on each plate. Petri plates were incubated at 37 C for 16 hr and then at room temperature for 2 to 7 days to allow maximum color development. Plates were screened visually with the aid of an illuminated 3x magnifier and a low power stage-dissecting microscope for the rare, pale pinkish-yellow to deep-yellow colonies that could be observed in the background of pink colonies. A colony color of yellow or pinkish-yellow was taken as presumptive evidence of a cyclization activity. These yellow colonies were collected with sterile toothpicks and used to inoculate 3ml of LB medium in culture tubes with overnight growth at 37°C and shaking at 225 cycles/min. Cultures were split into two aliquots in microfuge tubes and harvested by centrifugation at a setting of 5 in an Eppendorf 5415C microfuge. After discarding the liquid, one pellet was frozen for later purification of plasmid DNA. To the second pellet was added 1.5 ml EtOH, and the pellet was resuspended by vortex mixing, and extraction was allowed to proceed in the dark for 15-30 min with occasional remixing. Insoluble materials were pelleted by centrifugation at maximum speed for 10 min in a microfuge. Absorption spectra of the supernatant fluids were recorded from 350-550 nm with a Perkin Elmer lambda six spectrophotometer.

Analysis of isolated clones

20

5

10

15

Eight of the yellow colonies contained β -carotene indicating that a single gene product catalyzes both cyclizations required to form the two β endgroups of the symmetrical β -carotene from the symmetrical precursor lycopene. One of the yellow colonies contained a pigment with the spectrum characteristic of δ -carotene, a monocyclic carotenoid with a single ϵ endgroup. Unlike the β cyclase, this ϵ -cyclase appears unable to carry out a second cyclization at the other end of the molecule.

25

The observation that ϵ -cyclase is unable to form two cyclic ϵ -endgroups (e.g. the bicyclic ϵ -carotene) illuminates the mechanism by which plants can coordinate and control the flow of substrate into carotenoids derived from β -carotene versus those derived from α -carotene and also can prevent the formation of carotenoids with two ϵ endgroups.

30

The availability of the A. thaliana gene encoding the ϵ -cyclase enables the directed manipulation of plant and algal species for modification of carotenoid content and

composition. Through inactivation of the ϵ -cyclase, whether at the gene level by deletion of the gene or by insertional inactivation or by reduction of the amount of enzyme formed (by such as antisense technology), one may increase the formation of β -carotene and other pigments derived from it. Since vitamin A is derived only from carotenoids with β endgroups, an enhancement of the production of β -carotene versus α -carotene may enhance nutritional value of crop plants. Reduction of carotenoids with ϵ -endgroups may also be of value in modifying the color properties of crop plants and specific tissues of these plants. Alternatively, where production of α -carotene, or pigments such as lutein that are derived from α -carotene, is desirable, whether for the color properties, nutritional value or other reason, one may overexpress the ϵ -cyclase or express it in specific tissues. Wherever agronomic value of a crop is related to pigmentation provided by carotenoid pigments the directed manipulation of expression of the ϵ -cyclase gene and/or production of the enzyme may be of commercial value.

The predicted amino acid sequence of the A. thaliana ϵ -cyclase enzyme was determined. A comparison of the amino acid sequences of the β - and ϵ -cyclase enzymes of Arabidopsis thaliana (Fig. 13) as predicted by the DNA sequence of the respective cDNAs (Fig. 4 for the ϵ -cyclase cDNA sequence), indicates that these two enzymes have many regions of sequence similarity, but they are only about 37% identical overall at the amino acid level. The degree of sequence identity at the DNA base level, only about 50%, is sufficiently low such that we and others have been unable to detect this gene by hybridization using the β cyclase as a probe in DNA gel blot experiments.

REFERENCES

Each reference cited in this application and/or listed below is hereby incorporated by reference.

- 25 Bird et al, 1991 Biotechnology 9, 635-639.
 - Bishop et al., (1995) FEBS Lett. 367, 158-162.
 - Bramley, P.M. (1985) Adv. Lipid Res. 21, 243-279.
 - Bramley, P.M. (1992) Plant J. 2, 343-349.
- Britton, G. (1988). Biosynthesis of carotenoids. In Plant Pigments, T.W. Goodwin, ed. (London: Academic Press), pp. 133-182.

5

10

15

20

10

15

20

25

30

09/63055

Britton, G. (1979) Z. Naturforsch. Section C Biosci. 34, 979-985.

Britton, G. (1995) UV/Visible spectroscopy. In Carotenoids, Vol. IB: Spectroscopy,

G. Britton, S. Liaaen-Jensen, H.P. Pfander, eds. (Basel: Birkhauser Verlag), pp. 13-62.

Bouvier et al., (1994) Plant J. 6, 45-54.

Cunningham et al., (1985) Photochem. Photobiol. 42: 295-307.

Cunningham et al., (1993) FEBS Lett. 328, 130-138.

Cunningham et al., (1994) Plant Cell 6, 1107-1121.

Cunningham et al., (1996) Plant Cell 8, 1613-1626.

Davies, B.H. (1976). Carotenoids. In Chemistry and Biochemistry of Plant Pigments,

Vol. 2, T.W. Goodwin, ed (New York: Academic Press), pp. 38-165.

Del Sal et al., (1988). Nucl. Acids Res. 16, 9878.

Demmig-Adams & Adams, (1992) Ann. Rev. Plant Physiol. Mol. Biol. 43, 599-626.

Enzell & Back, (1995) Mass spectrometry. In Carotenoids, Vol. IB: Spectroscopy, G.

Britton, S. Liaaen-Jensen, H.P. Pfander, eds. (Basel: Birkhauser Verlag), pp. 261-320.

Frank & Cogdell (1993) Photochemistry and function of carotenoids in photosynthesis. In Carotenoids in Photosynthesis. A. Young and G. Britton, eds. (London: Chapman and Hall). pp. 253-326.

Goodwin, T.W. (1980). The Biochemistry of the Carotenoids. 2nd ed, Vol. 1 (London: Chapman and Hall.

Horvath et al., (1972) Phytochem. 11, 183-187.

Hugueney et al., (1995) Plant J. 8, 417-424.

Hundle et al., (1991) Photochem. Photobiol. 54, 89-93.

Jensen & Jensen, (1971) Methods Enzymol. 23, 586-602.

Kargl & Quackenbush, (1960) Archives Biochem. Biophys. 88, 59-63.

Kargl et al., (1960) Proc. Am. Hort. Soc. 75, 574-578.

Kieber et al., (1993) Cell 72, 427-441.

Koyama, Y. (1991) J. Photochem. Photobiol., B, 9, 265-80.

Krinsky, N.I. (1987) Medical uses of carotenoids. In Carotenoids, N.I. Krinsky, M.M. Mathews-Roth, and R.F. Taylor, eds. (New York: Plenum), pp. 195-206.

Kyte & Doolittle, (1982) J. Mol. Biol. 157, 105-132.

LaRossa & Schloss, (1984) J. Biol. Chem. 259, 8753-8757.

Misawa et al., (1994a) Plant J. 6, 481-489.

-

Misawa et al., (1994b) J. Biochem, Tokyo, 116, 980-985.

Norris et al., (1995) Plant Cell 7, 2139-2149.

Pecker et al., (1996) Submitted to Plant Mol. Biol.

Perry et al., (1986) J. Bacteriol. 168, 607-612.

Persson & Argos, (1994) J. Mol. Biol. 237, 182-192.

Plumley & Schmidt, (1987) Proc. Nat. Acad. Sci. USA 83, 146-150.

Plumley & Schmidt, (1995) Plant Cell 7, 689-704.

Rossmann et al., (1974) Nature 250, 194-199.

Rock & Zeevaart (1991) Proc. Nat. Acad. Sci. USA 88, 7496-7499.

10 Rost et al., (1995) Protein Science 4, 521-533.

Sambrook et al., (1989) Molecular Cloning: A Laboratory Manual, 2nd edition (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Sancar, A. (1994) Biochemistry 33, 2-9.

Sander & Schneider, (1991) Proteins 9, 56-68.

15 Sandmann, G. (1994) Eur. J. Biochem. 223, 7-24.

Scolnik & Bartley, (1995) Plant Physiol. 108, 1342.

Siefermann-Harms, D. (1987) Physiol. Plant. 69, 561-568.

Spurgeon & Porter, (1980). Biosynthesis of carotenoids. In Biochemistry of

Isoprenoid Compounds, J.W. Porter, and S.L. Spurgeon, eds. (New York: Wiley), pp. 1-122.

20 Tomes, M.L. (1963) Bot. Gaz. 124, 180-185.

Tomes, M.L. (1967) Genetics 56, 227-232.

Tuveson et al., (1986) J. Bacteriol. 170, 4675-4680.

Van Beeumen et al., (1991) J. Biol. Chem. 266, 12921-12931.

Weedon & Moss, (1995) Structure and Nomenclature. In Carotenoids, Vol. IB:

25 Spectroscopy, G. Britton, S. Liaaen-Jensen, H.P. Pfander, eds. (Basel: Birkhauser Verlag), pp. 27-70.

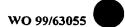
Wierenga et al., (1986) J. Mol. Biol. 187, 101-107.

Zechmeister, L. (1962) Cis-Trans Isomeric Carotenoids, Vitamins A and Arylpolyenes. Springer-Verlag, Vienna.

Having now fully described the invention, it will be apparent to one of ordinary skill in the art that many changes and modifications can be made thereto without departing from the spirit or scope of the invention as set forth herein.

We claim:

- 1. An isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase enzyme activity and has an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23 or 25-27.
- 5 2. The nucleic acid sequence of claim 1, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23 or 25-27.
 - 3. A vector comprising the nucleic acid sequence of claim 1, wherein the nucleic acid sequence is operably linked to a promoter.
 - 4. A host cell which contains the vector of claim 3.
- 5. The host cell of claim 4, wherein the host cell is selected from the group consisting of a bacterial cell, an algal cell, a yeast cell and a plant cell.
 - 6. The host cell of claim 4, wherein the host cell is a photosynthetic cell.
- 7. An isolated and/or purified protein having lycopene ε-cyclase enzyme activity and having an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23 or
 25-27.
 - 8. The protein of claim 7, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23 or 25-27.



AMENDED CLAIMS

[received by the International Bureau on 15 November 1999 (15.11.99); original claims 1,2,7 and 8 amended; remaining claims unchanged (1 page)]

- 1. An isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase enzyme activity and has an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23, 25 or 26.
- 5 2. The nucleic acid sequence of claim 1, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23, 25 or 26.
 - 3. A vector comprising the nucleic acid sequence of claim 1, wherein the nucleic acid sequence is operably linked to a promoter.
 - 4. A host cell which contains the vector of claim 3.
- 5. The host cell of claim 4, wherein the host cell is selected from the group consisting of a bacterial cell, an algal cell, a yeast cell and a plant cell.
 - 6. The host cell of claim 4, wherein the host cell is a photosynthetic cell.
- 7. An isolated and/or purified protein having lycopene ε-cyclase enzyme activity and having an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23, 25 or 26.
 - 8. The protein of claim 7, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23, 25 or 26.

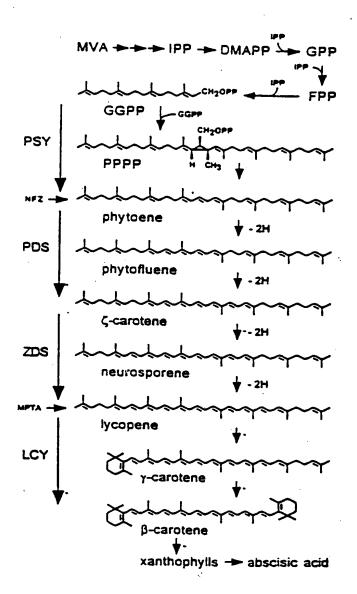


FIGURE 1

£.

FIGURE 2

Arabidopsis thaliana epsilon cyclase:

FIGURE 4

getettete etecteetet accgatttee gaeteegeet eccgaaatee ttatccggat teteteegte tettegattt aaacgetttt etgtetgtta 51 cgtcgtcgaa gaacggagac agaattctcc gattgagaac gatgagagac 101 eggagageae gageteeaea aacgetatag aegetgagta tetggegttg 151 cgtttggcgg agaaattgga gaggaagaaa tcggagaggt ccacttatct 201 aatcgctgct atgttgtcga gctttggtat cacttctatg gctgttatgg 251 ctgtttacta cagattetet tggcaaatgg agggaggtga gatetcaatg 301 ttggaaatgt ttggtacatt tgctctctct gttggtgctg ctgttggtat 351 ggaattotgg gcaagatggg otcatagage totgtggcac gottototat 401 ggaatatgca tgagtcacat cacaaaccaa gagaaggacc gtttgagcta 451 501 aacgatgttt ttgctatagt gaacgetggt ecagegattg gteteetete ttatggattc ttcaataaag gactcgttcc tggtctctgc tttggcgccg 551 ggttaggcat aacggtgttt ggaatcgcct acatgtttgt ccacgatggt 601 ctogtgcaca agogtttooc tgtaggtooc atogoogacg tooottacot 651 ccgaaaggte geogeogete accagetaca teacacagae aagtteaatg 701 gtgtaccata tggactgttt cttggaccca aggaattgga agaagttgga 751 ggaaatgaag agttagataa ggagattagt cggagaatca aatcatacaa 801 aaaggeeteg ggeteegggt egagttegag ttettgaett taaacaagtt 851 ttaaatccca aattcttttt ttgtcttctg tcattatgat catcttaaga 901 951 cggtct

FIGURE 5

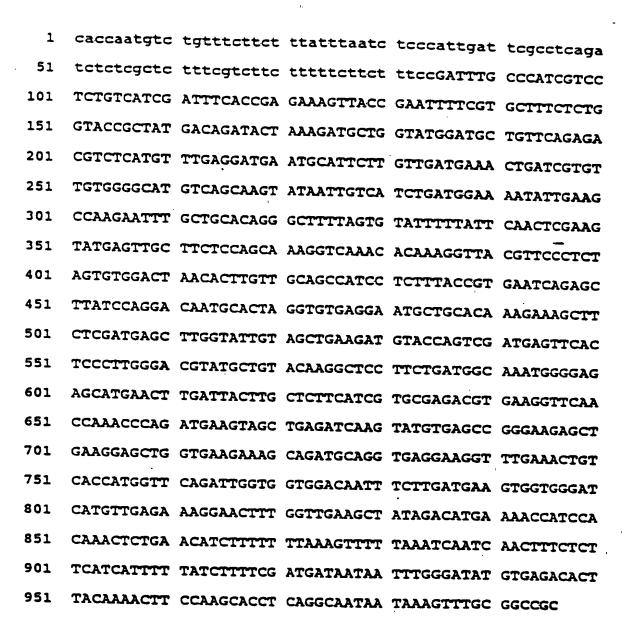
<u>`</u>

| A. chal. | | 884 6 | SSTORLLLE | Kslegopspse | RIMAISVOTV | 44) Spss sstopalald ksisgpspsu rphopsycy verropsps addresses stallalary. | ENDER LETS | 49 TANGINALS |
|---|--|--|--|--|--|---|--|---|
| A. thal. Alical. A. auranc. E. berb. E. ured. Consensus | MALAETA EN | KKSERSTYLI | AMESSIGIT SHAWAYTA | SHAVPAVTTR | MOAS | SPGENFOTFA HTOFL HTNFL HG. HG. | LSVGAAVGHI IVVATVLVHI IVVATVLVHI IVILSVIAHI IVFVTVIGHI | LTAYSVORMI LTAYSVORMI LTAYSVORMI GLAFTFORYI VIALANITYI ANE |
| | | | Predicted TM helix | helix | | Predic | Predicted TM helix | |
| A. chal. Alical. | WHASE, WROTH MHGPLGWOM | | PELMOVPALV LEXMOLYGVV | Nagpaigles Favlatilet | YGPPHKOLVP VGAYNMPVLM | GLCFGAGLGI HIALCH | TVPGLATHEV TVPGLIYFIL | TLL HDGLVHXR.PP HDGLVHQR.MP |
| F. herb. F. ured. Consensus | HORGENCHAM HORGENCHAM HORGENCHAM HORGENCHAM HORGENCHAM | KSHRIEDOHA ESHRIPRKOV LSHRIPRKOA | LECHOLIGE PATATALTA FELHOLIAVY PACVALALIA FEVNOLYAVY PALSILLIY | PACATALIA | VGTAGVNPLQ LGSTGMNPLQ | WI GAGA | TWGLLYFW TWGLLYFW TAYGLLYFW | HDGLVHQRWP HDGLVHQRWP HDGLVHQRWP |
| | | | Pre | Predicted TM heix | , | Predicted TM helix | M helix | KUSLVN-R-P |
| A. thal. Alical. A. surant. S. berb. B. ured. Consensus | VGPLADVPTL FRYIPRGYT FRYIPRGYA FHWIPRRGYL FRYIPRKGYL | RECYGANGLH RELYGANGLH RELYGANGLH RELYGANGLH RELYVANGLH RELYVANGLH | HT. DKTMGV HAVEGROHCV HAVBGREDCV HAVRGREGCV HAVRGKEGCV | PYGLFLGPKE SPGFTYAPP. SPGFTYAR. SPGFTYAR. | LEFVOGNEEL VDICKOLICH VDICKOLICH PADICATIRE LSTICATIRE | DICEI SRRIKS SGVLRPQDER SGVLRAEAGE RHGRPPICEDA RHGARAGA | YXXASGSGSS PS* RT* AXDREDAASP ARDAGGGEDE | 506 555° 55555PE* |

7 / 30

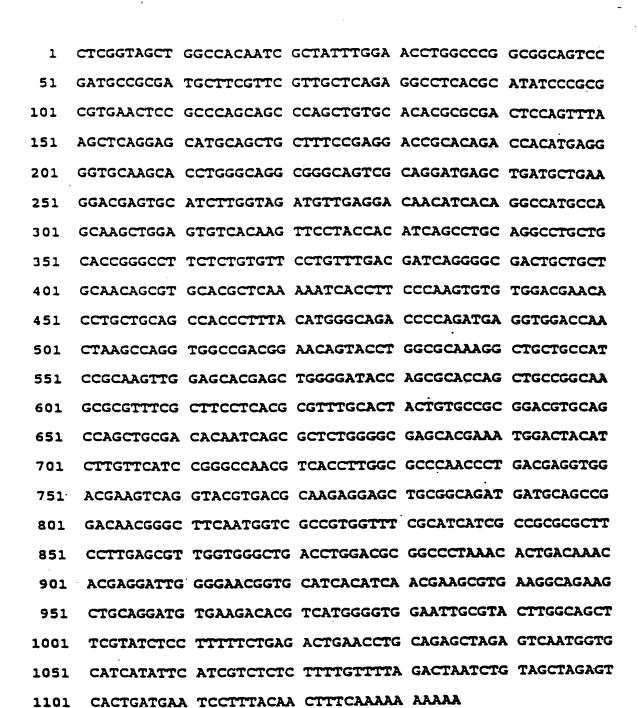
ccacgggtcc gcctccccgt ttttttccga tccgatctcc ggtgccgagg 1 actcagctgt ttgttcgcgc tttctcagcc gtcaccatga ccgattctaa 51 101 cgatgctgga atggatgctg ttcagagacg actcatgttt gaagacgaat gcattctcgt tgatgaaaat aatcgtgtgg tgggacatga cactaagtat 151 201 aactgtcatc tgatggaaaa gattgaagct gagaatttac ttcacagagc tttcagtgtg tttttattca actccaagta tgagttgctt ctccagcaac 251 301 ggtcaaaaac aaaggttact ttcccacttg tgtggacaaa cacttgttgc 351 agccatcctc tttaccgtga atccgagctt attgaagaga atgtgcttgg 401 tgtaagaaat gccgcacaaa ggaagctttt cgatgagctc ggtattgtag 451 cagaagatgt accagtcgat gagttcactc ccttgggacg catgctttac 501 aaggcacett etgatgggaa atggggagag cacgaagttg actatetact 551 cttcatcgtg cgggatgtga agcttcaacc aaacccagat gaagtggctg 601 agatcaagta cgtgagcagg gaagagctta aggagctggt gaagaaagca 651 gatgctggcg atgaagctgt gaaactatct ccatggttca gattggtggt ggataatttc ttgatgaagt ggtgggatca tgttgagaaa ggaactatca 701 ctgaagctgc agacatgaaa accattcaca agctctgaac tttccataag 751 ttttggatct tccccttccc ataataaaat taagagatga gacttttatt 801 gattacagac aaaactggca acaaaatcta ttcctaggat ttttttttgc 851 tttttattta cttttgattc atctctagtt tagttttcat cttaaaaaaa 901 951 aaaa





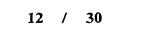


CTCGGTAGCT GGCCACAATC GCTATTTGGA ACCTGGCCCG GCGGCAGTCC GATGCCGCGA TGCTTCGTTC GTTGCTCAGA GGCCTCACGC ATATCCCCCG CGTGAACTCC GCCCAGCAGC CCAGCTGTGC ACACGCGCGA CTCCAGTTTA 101 AGCTCAGGAG CATGCAGATG ACGCTCATGC AGCCCAGCAT CTCAGCCAAT 151 CTGTCGCGCG CCGAGGACCG CACAGACCAC ATGAGGGGTG CAAGCACCTG 201 251 GGCAGGCGGG CAGTCGCAGG ATGAGCTGAT GCTGAAGGAC GAGTGCATCT 301 TGGTGGATGT TGAGGACAAC ATCACAGGCC ATGCCAGCAA GCTGGAGTGT 351 CACAAGTTCC TACCACATCA GCCTGCAGGC CTGCTGCACC GGGCCTTCTC 401 TGTGTTCCTG TTTGACGATC AGGGGCGACT GCTGCTGCAA CAGCGTGCAC GCTCAAAAAT CACCTTCCCA AGTGTGTGGA CGAACACCTG CTGCAGCCAC 451 501 CCTTTACATG GGCAGACCCC AGATGAGGTG GACCAACTAA GCCAGGTGGC CGACGGAACA GTACCTGGCG CAAAGGCTGC TGCCATCCGC AAGTTGGAGC 551 ACGAGCTGGG GATACCAGCG CACCAGCTGC CGGCAAGCGC GTTTCGCTTC 601 CTCACGCGTT TGCACTACTG TGCCGCGGAC GTGCAGCCAG CTGCGACACA ATCAGCGCTC TGGGGCGAGC ACGAAATGGA CTACATCTTG TTCATCCGGG 701 CCAACGTCAC CTTGGCGCCC AACCCTGACG AGGTGGACGA AGTCAGGTAC 751 GTGACGCAAG AGGAGCTGCG GCAGATGATG CAGCCGGACA ACGGGCTGCA 801 ATGGTCGCCG TGGTTTCGCA TCATCGCCGC GCGCTTCCTT GAGCGTTGGT 851 GGGCTGACCT GGACGCGGCC CTAAACACTG ACAAACACGA GGATTGGGGA 901 ACGGTGCATC ACATCAACGA AGCGTGAAAG CAGAAGCTGC AGGATGTGAA 951 GACACGTCAT GGGGTGGAAT TGCGTACTTG GCAGCTTCGT ATCTCCTTTT 1001 TCTGAGACTG AACCTGCAGT CAGGTCCCAC AAGGTCAGGT AAAATGGCTC GATAAAATGT ACCGTCACTT TTTGTCGCGT ATACTGAACT CCAAGAGGTC 1101 ΑΑΑΑΑ ΑΑΑΑΑΑΑ 1151



| HP04 | 1 MLRSLLRGLT | HTPRVNSAGO | DECAMADI OR | VI Dayones | 50 |
|----------|--------------------------|--------------------------|--------------------------|--------------------------|------------|
| HP05 | MLRSLLRGLT | HIPRVNSAQQ | PSCAHARLOF | KLRSMOMTLM | |
| ATDP7 | MSVSSLFNLP | LIRLRSLA. | LSSSFSSFRF | KLRSMQLL AHRPLSSIS. | DDWID |
| c brew. | MS.SSMLNFT | .ASRIVSLPL | LSSPPSRVHL | PLCFFSPISL | PRKLPNFRAF |
| ATOPS | • • • • • • • • • | .TGPPPRFFP | IRSPVPRTOL | FVRAFSAV | TORFSAKLTF |
| S ceres. | MTADNNSM | PHGAVSSYAK | LVQNQTPEDI | LEEFPEIIPL | QQRPN TR |
| | | | | | 22.00.002 |
| | 51 | | | | 100 |
| | AEDRTDHMRG SEDRTDHMRG | ASTWAGGQSQ | DELMLKDECI | LVDVEDNITG | HASKLECHKF |
| | SGTA.MTD | ASTWAGGQSQ TKDAGMDAVQ | DELMLKDECI | LVDVEDNITG | HASKLECHKF |
| | SSQATT.MGE | VVDAGMDAVQ | RRLMFEDECI RRLMFEDECI | LVDETDRVVG | HVSKYNCHLM |
| | | SNDAGMDAVQ | RRLMFEDECI | LVDENDKVVG LVDENNRVVG | HESKYNCHLM |
| | SSETSNDESG | | QIKLMNENCI | VLDWDDNAIG | |
| | | | | ADMODRATG | AGIKKACHIM |
| | 101 | | | | 150 |
| | LPHQPAGLLH | RAFSVFLFDD | QGRLLLQQRA | RSKITFPSVW | TNTCCSHPLH |
| | LPHQPAGLLH | RAFSVFLFDD | QGRLLLQQRA | RSKITFPSVW | TNTCCSHPLH |
| | ENIEAKNLLH | RAFSVFLFNS | KYELLLQQRS | NTKVTFPLVW | TNTCCSHPLY |
| | EKIESENLLH | RAFSVFLFNS | KYELLLQQRS | ATKVTFPLVW | TNTCCSHPLY |
| | EKIEAENLLH | RAFSVELENS | KYELLLQQRS | KTKVTFPLVW | TNTCCSHPLY |
| | LATE. NGLLIA | KWLZALTINE | QGELLLLQQRA | TEXITFPDLW | TNTCCSHPLC |
| | 151 | | | | 200 |
| | GQTPDEVDQL | SQVADGTVPG | AKAAAIRKLE | HELGIPAHQL | PA SAFPEIT |
| | GQTPDEVDQL | SQVADGTVPG | | HELGIPAHQL | PA.SAFRELT |
| | RE | | VRNAAQRKLL | DELGIVAEDV | PV.DEFTPLG |
| | RE | | VRNAAQRKLL | DELGIPAEDL | PV.DQFIPLS |
| | RE | | VRNAAQRKLF | DELGIVAEDV | PV.DEFTPLG |
| | IDDELGL | KGKLDDKIKG | AITAAVRKLD | HELGIPEDET | KTRGKFHFLN |
| | 201 | | | 1 | |
| | RLHYCAADVQ | PAATQSALWG | EHEMDYILFI | RANVTL | 250 |
| | RLHYCAADVQ | | EHEMDYILFI | RANVIL | |
| | RMLY | . KAPSDGKWG | EHELDYLLFI | VRDVKV | |
| | RILY | .KAPSDGKWG | EHELDYLLFI | | DPNPDEVAEV |
| | RMLY | . Kapsdgkwg | EHEVDYLLFI | VRDVKL | |
| | RIHY | .MAPSNEPWG | EHEIDYILFY | KINAKENLTV | NPNVNEVRDF |
| | 261 | | | | |
| | 251 | MMO DOM | CT ALICATION T | | 300 |
| | RYVTOFFLEO | MMO PDN | GLQWSPWFRI | IAARFLERWW IAARFLERWW | ADLDAALNTD |
| | KYVSREELKE | LVKKADAGEE | GLUMSPALKT | VVDNFLMKWW | AULDAALNTD |
| | KYMNRDDLKE | LLRKADAEEE | GUKLSPWFRL | VVDNFLFKWW | DHVEKGILVE |
| | KYVSREELKE | LVKKADAGDE | AVKLSPWFRL | VVDNFLMKWW | DHVEKGTITE |
| | KWVSPNDLKT | MFADP | SYKFTPWFKI | ICENYLFNWW | EQLDDLSEVE |
| | | | | | |
| | 301 | T1777 : | | | |
| | KHEDWGTVHH | | | | • |
| | KHEDWGTVHH A.IDMKTIHK | | | | |
| | A. ADMKTIHK | | | | |
| | A. ADMKTIHK | | | • | |
| | NDRQIHR | | t :: | | |
| | | | • | | |

-1



WO 99/63055

ccaaaaacaa ctcaaatctc ctccgtcgct cttactccgc catgggtgac gactccggca tggatgctgt tcagcgacgt ctcatgtttg acgatgaatg 51 cattttggtg gatgagtgtg acaatgtggt gggacatgat accaaataca 101 attgtcactt gatggagaag attgaaacag gtaaaatgct gcacagagca 151 201 ttcagcgttt ttctattcaa ttcaaaatac gagttacttc ttcagcaacg 251 gtctgcaacc aaggtgacat ttcctttagt atggaccaac acctgttgca 301 gecatecaet etacagagaa teegagettg tteeegaaac geetgagaga 351 401 451 501 551 601 651 xxxxxxxxx xxxxxxxx xxxxxxxxx tcatgtgcaa aagggtacac 701 tcactgaatg caatttgata tgaaaaccat acacaagctg atatagaaac acacecteaa eegaaaagea ageetaataa ttegggttgg gtegggteta 751 ccatcaattg ttttttttt ttaacaactt ttaatctcta tttgagcatg 801 851 ttgattcttg tcttttgtgt gtaagatttt gggtttcgtt tcagttgtaa taatgaacca ttgatggttt gcaatttcaa gttcctatcg acatgtagtg 901 951 atctaaaaaa

FIGURE 12

J- __

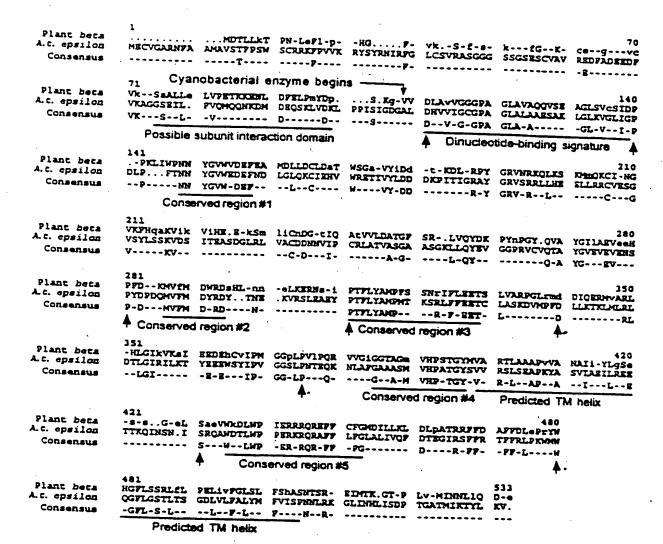


FIGURE 13





Adonis palaestina ε-cyclase cDNA #5 Length: 1898

| 1 | | | | | |
|--------|------------|--------------|--------------|------------|---------------|
| 51 | aaaggagtg | | | attettgea. | a cacttata |
| 101 | caaactcca | | tctcttcaaa | acaacaaac | |
| 151 | gagtatctg | | acttggtgtt | cgcaacctc | ,-,-,- |
| 201 | ccctgtgtg | g acttttggaa | a caagaaacct | | |
| 251 | ataacataca | ı togatatggt | tcttcttgta | | |
| 301 | gctgatggtg | | , tagaagttct | gttgcttata | J - 3 - 3 - 3 |
| 351 | tgtggatgaa | gaggatttt | tcaaagctg | tggttctgag | |
| _ | tccaaatgca | gcaaacaaac | , tctatggaga | aacaggccaa | |
| 401 | aagttgccac | caatacctt | tggagaatco | gtgatggact | |
| 451 | aggttgtgga | cctgctggtc | tttcactggc | tgcagaagct | |
| 501 | ggttgaaagt | tggccttatt | ggtcctgatc | ttccttttac | |
| 551 | ggtgtgtggg | aagacgagtt | | ggacttgaac | |
| 601 | gcatgcttgg | aaggacacca | | | |
| 651 | ttattggtcg | tgcatatgga | cgagttagtc | | |
| 701 | ttgctgaaaa | ggtgtgtgga | gtcaggtgta | | 7-33-9 |
| 751 | ggaaaggatc | actgaagctg | gtgatggcca | | |
| 801 | atgagatctt | tatecettge | aggettgeta | | |
| 851 | tcagggaaac | ttttggagta | tgaagtaggt | | |
| 901 | aaccgcttat | ggggtggagg | ttgaggtgga | ggccctcgtg | |
| 951 | acttaatggt | attcatggac | tacagagact | | |
| 1001 | tgctcggaag | aagaatatcc | aacatttctC | atatgcaaca | |
| 1051 | aacaagactt | ttttttgagg | aaacctgttt | | |
| 1101 | cattcgatct | actgaagaga | aaactgatgt | ggcctcaaaa | gatgccatgc |
| 1151 | atccaagtta | caaaagttta | tgaagaggaa | cacgattgaa | gactctgggt |
| 1201 | tggttcttta | ccaaacacag | agcaaaagaa | tggtcatata | ttcctgttgg |
| 1251 | caagcatggt | gcatccagca | acaggctatt | cctagcattt | ggtgctgcag |
| 1301 | gaagctccaa | aatatgcttc | tgtaattgca | cggttgtacg | gtcactgtca |
| 1351 | ctctgcgtat | gtggtttctg | gacaaagtag | aagattttga | agcaagataa |
| 1401 | aagcatggag | cagtctttgg | ccaaaggagc | tgcagtaaac | atttcaatgc |
| 1451 | tttctttttg | gattagaget | tattgtgcag | gaaaacgtca | aagagcatTc |
| 1501 | aacattcttt | agaaccttct | teegettgee | ctagatattg | aagcaaccag |
| 1551 | tccttgggtc | ttcactatca | tctttcgatc | aacttggatg | tggtggggtt |
| 1601 | atgtttgttt | tggcgccaaa | cagcatgagg | tcgtcttgtt | ttccatgtac |
| 1651 | gctttcagat | ccttctggtg | cagttatggt | atgtcacttg | tgagacattt |
| 1701 . | agtctcatct | attattaaac | tctagtgttt | aagagcttac | ctcgaaaggt |
| 1751 | ttcgaatgtg | tatatgatca | tctctatgta | caccaaataa | atgaggatcc |
| 1801 | taaagtaaat | gccgggtttg | | tatcctgtac | tctaatctca |
| 1851 | aaagtaaatt | tattgataca | | gtcaaaccgg | ccaatgatat |
| | = | - 3 =04 | 9-04-66 | ttttccttaa | aaaaaaa |

Adonis palaestina ε-cyclase #5 predicted polypeptide TRANSLATE from: 113 to: 1702 Length: 529 amino acids

| 351 KVYEEEWSYI PVGGSLPNTE QKNLAFGAAA SMVHPATGYS VVRSLSEA 401 YASVIAKILK QDNSAYVVSG QSSAVNISMQ AWSSLWPKER KRQRAFFL 451 LELIVQLDIE ATRTFFRTFF RLPTWMWWGF LGSSLSSFDL VLFSMYMF 501 APNSMRMSLV RHLLSDPSGA VMVRAYLER* | 201 251 301 351 401 451 | CVESGVSYLD LEYEVGGPRV EYPTFLYVMP KVYEEEWSYI YASVIAKILK LELIVQLDIE | VVIGCGPAGL CIEHAWKDTI SKVERITEAG CVQTAYGVEV MSPTRLFFEE PVGGSLPNTE QDNSAYVVSG ATRTFFRTFF | DGHSLVVCEN EVENNPYDPN TCLASKDAMP QKNLAFGAAA QSSAVNISMQ RLPTWMWWGF | IGRAYGRVSR EIFIPCRLAT LMVFMDYRDY FDLLKRKLMS SMVHPATGYS AWSSLWPKER | PFTNNYGVW HLLHEELLK VASGAASGK MQQKLQCSE RLKTLGIQV VVRSLSEAP KRORAFFI E |
|--|--|--|--|--|--|--|
|--|--|--|--|--|--|--|



301

351

DNA sequence of pot 3 cDNA (GenBank R27545) obtained from Nicholas J. Provart potato.seq Length: 1378 August 2, 1996 13:06 Type: N Check: 605 1 tageggnnnn naggatgagt teaaagatet tggtetteaa geetgeattg 51 aacatgtttg gcgggatacc attgtatatc ttgatgatga tgatcctatt 101 cttattggcc gtgcctatgg aagagttagt cgccatttac tgcacgagga 151 gttactcaaa aggtgtgtgg aggcaggtgt tttgtatcta aactcgaaag 201 tggataggat tgttgaggcc acaaatggcc acagtcttgt agagtgcgag 251 ggtgatgttg tgattccctg caggtttgtg actgttgcat cgggagcagc 301 ctcggggaaa ttcttgcagt atgagttggg aggtcctaga gtttctgttc 351 aaacagetta tggagtggaa gttgaggteg ataacaatee atttgaeeeg 401 ageetgatgg tttteatgga ttatagagae tatgteagae acgaegetea 451 atetttagaa getaaatate caacatttet etatgeeatg eccatgtete 501 caacacgagt ctttttcgag gaaacttgtt tggcttcaaa agatgcaatg 551 ccattcgatc tgttaaagaa aaaattgatg ttacgattga acaccetcgg
601 tgtaaagaatt aaagaaattt atgaggagga atggtcttac ataccagttg
651 gaggatcttt gccaaataca gaacaaaaaa cacttgcatt tggtgctgct 701 gctagcatgg ttcatccagc cacaggttat tcagtcgtca gatcactgtc 751 tgaagctcca aaatgcgcct tcgtgcttgc aaatatatta cgacaaaatc 801 atagcaagaa tatgcttact agttcaagta ccccgagtat ttcaactcaa 851 gettggaaca etetttggee acaagaacga aaacgacaaa gategttttt 901 cctatttgga ctggctctga tattgcagct ggatattgag gggataaggt 951 cattetecg egegteette egtgtgccaa aatggatgtg geagggattt 1001 ettggttcaa gtetteettn ageagacete atgttatttg cettetacat 1051 gtttattatt gcaccaaatg acatgagaag aggcttaatc agacatcttt 1101 tatctgatcc tactggtgca acattgataa gaacttatct tacattttag 1151 agtaaattcc tcctacaata gttgttgaan nagaggcctc attacttcag 1201 attcataaca gaaatcgcgg tctctcgagg ccttgtatat aacattttca 1251 ctaggttaat attgcttgaa taagttgcac agtttcagtt tttgtatctg 1301 cttcttttt gtccaagatc atgtattgan ccaatttata tacattgcca 1351 gratatataa attttataaa aaaaaaaa TRANSLATE from: 14 to: 1147 poteps.pep Length: 378 1 DEFKDLGLQA CIEHVWRDTI VYLDDDDPIL IGRAYGRVSR HLLHEELLKR CVEAGVLYLN SKVDRIVEAT NGHSLVECEG DVVIPCRFVT VASGAASGKF 51 LQYELGGPRV SVQTAYGVEV EVDNNPFDPS LMVFMDYRDY VRHDAQSLEA 101 151 KYPTFLYAMP MSPTRVFFEE TCLASKDAMP FDLLKKKIML RINTLGVRIK 201 EIYEEEWSYI PVGGSLPNTE QKTLAFGAAA SMVHPATGYS VVRSLSEAPK CAFVLANILR QNHSKNMLTS SSTPSISTQA WNTLWPQERK RQRSFFLFGL 251

FIGURE 15A

ALILQLDIEG IRSFFRAFFR VPKWMWQGFL GSSLSXADLM LFAFYMFIIA

PNDMRRGLIR HLLSDPTGAT LIRTYLTF*

Chimeric lettuce/potato lycopene ε-cyclase: converts lycopene to δ-carotene, the lettuce cDNA converts lycopene to ε-carotene and the potato cDNA does not produce an active enzyme

(amino acids in lower case are from lettuce and those in uppercase are from the potato cDNA; an AvaII site in common to the two cDNAs was used to construct the chimera)

```
1 mecfgarnmt atmavftcpr ftdcnirhkf sllkqrftn lsassslrqi
51 kcsaksdrcv vdkqqisvad eedyvkagqs
101 eklaqipiqn cildlvviqc qpaqlalaae saklqlnvql iqpdlpftnn
151 yqvwqdefiq lqleqciehs wkdtlvyldd adpiriqray qrvhrdllhe
201 ellrrcvesq vsylsskver iteapnqysl iecegnitip crlatvasqa
251 asqkfleyel qGPRVSVQTA YGVEVEVDNN PFDPSLMVFM DYRDYVRHDA
101 QSLEAKYPTF LYAMPMSPTR VFFEETCLAS KDAMPFDLLK KKLMLRLNTL
151 GVRIKEIYEE EWSYIPVGGS LPNTEQKTLA FGAAASMVHP ATGYSVVRSL
152 GVRIKEIYEE EWSYIPVGGS LPNTEQKTLA FGAAASMVHP PQERKRQRSF
153 GVRIKEIYEE EWSYIPVGGS LPNTEQKTLA FGAAASMVHP PQERKRQRSF
154 FLFGLALILQ LDIEGIRSFF RAFFRVPKWM WQGFLGSSLS XADLMLFAFY
150 MFIIAPNDMR RGLIRHLLSD PTGATLIRTY LTF*
```

FIGURE 15B



| GAP comparison of Arabidopsis e-cyclase x potato e-cyclase (partial) |
|---|
| blosum62.cmp Gap Weight: 12 Average Match: 2.912 Length Weight: 4 Average Mismatch: -2.003 Quality: 1485 Ratio: 3.929 Percent Similarity: 79.893 Percent Identity: 76.139 Match display thresholds for the alignment(s): 1 Torongan |
| Quality: 1485 Length: 529 Ratio: 3.929 Gaps: 1 |
| fatch display thresholds for the alignment(s): = IDENTITY : = 2 . = 1 |
| 151 EDEFNDLGLOKCIEHVWRETIVYLDDDKPITIGRAYGRVSRRLLHEELLR 200 |
| 201 RCVESGVSYLSSKVDSITEASDGLRLVACDDNNVIPCRLATVASGAASGK 250 |
| 251 LLQYEVGGPRVCVQTAYGVEVEVENSPYDPDQMVFMDYRDYTNEKVRSLE 300 |
| THE TAX TO |
| 301 AEYPTFLYAMPMTKSRLFFEETCLASKDVMPFDLLKTKLMLRLDTLGIRI 350 |
| 351 LKTYEEEWSYIPVGGSI PNTFOYNI A SCANA CHANGA |
| . |
| 401 KYASVIAFILBETTYOI |
| |
| 446 LALIVOFDIEGIRSFERIEFRI PERPENDICE |
| |
| 496 SPNNLRKGLINHLISDPTGATMIKTYLKV 524 |
| 350 APNDMRRGLIRHLISDPTGATI IPTVITE 270 |

FIGURE 16



| Adonis p | <i>alaes</i> | tina | Ipil |
|----------|--------------|------|------|
|----------|--------------|------|------|

| _ | - | | | | |
|--------|------------|------------|-------------|------------|------------|
| 1. | attcatcttc | agcagcgctg | tcgtactctt | tctatatctt | cttccatcac |
| 51 | taacagtagt | cgccgacggt | tgaatcggct | attcgcctca | acgtcaacta |
| 101 | | cactgatgct | | ctgttcagaa | gcggctcatg |
| 151 | | aatgtatttt | | aatgacaagg | tegtegggea |
| 201 | tgattccaaa | tacaactgtc | atttgatgga | aaagatagag | |
| 251 | tgcttcacag | agccttcagt | attttcttat | tcaactcaaa | |
| 301 | cttcttcage | aacgatccgc | cacaaaggta | acattcccc | atatgaattg |
| 351 | aaacacatot | tgcagtcatc | ctctctttca | teatteear | tcgtatggac |
| 401 | aaaattatct | cggtgtacga | aacoctocac | aaaaaaaaa | ctcatagaag |
| 451 | Ctaggcattc | cagctgaaga | tataccaatt | gatgaatet. | tttagacgag |
| 501 | togcattorr | tacaaagctc | catctgacga | gatyaatta | |
| 551 | togactator | cctatttatt | caccegacgg, | Caaacgggga | gagcacgaat |
| 601 | Catcaacttc | cctatttatt | gracyayary | Lyaaatacga | |
| 651 | actgaageeg | ctgatgctaa | gtatgttaat | cgcgaggagt | tgagagagat |
| 701 | ttagattagt | gctgatgctg | gcgaagaggg | actcaagttg | |
| 751 | cagattygt | tgttgataac | tttttgttca | agtggtggga | tcatgtagag |
| 801 | ttaagaaaaa | ttaaggaagt | tgctgacatg | aaaactatcc | acaagttgac |
| 851 | ctaayaggac | ttctctcctc | tgttctacta | tttgttttt | gctacaataa |
| 901 | gragarage | ataagcagtt | tttctgtttt | ctttaattta | tggcttttga |
| 951 | atttgcctcg | atgttgaact | tgtaacatat | ttagacaaat | atgagacett |
| 1001 | gtaagttgaa | tttgaggctg | aatttatatt | tttgggaaca | taataatgtt |
| 1 (11) | 22 | | | | • |

FIGURE 17A

Adonis palaestina Ipi2

| pa. | reescria ip | 12 | | | |
|---------|-------------|------------|-------------|-------------------|------------|
| 1 | ttttaaagct | ctttcgctcc | accaccatca | aagccagcca | aatttctctg |
| 51 | uncauaget. | aaaaacaccq | CTTTGGGGCTT | taacccctcc | 34345555 |
| 101 | ccttgtttac | gatacqcatc | taaaccagta | attetegett | ttaatttgtt |
| 151 | tcctaaatta | ggcccctttc | cogaatccc | 30334434 | cgtcgatcag |
| 201 | gattaatcct | ttatatagta | tettetees | Caccacacac | acattatcag |
| 251 | cttcqtqttc | ttctcccact | Ottcatcttc | Caccaccaaa | tegtactett |
| 301 | tctatttctt | Cttccatcac | taacagtcct | agcagegeeg | togtactett |
| 351 | gttcgcctca | acgtcgacta | tgggtgaagt | cyccgagggt | tgaatcggct |
| 401 | CCGtccagaa | acaacttata | ttccaccata | cyctgatget | ggtatggatg |
| 451 | aatgacaagg | tratragaca | ttcgacgatg | aatgtatttt | ggtggatgag |
| 501 | aaagatagag | acadaaaact | tgattccaaa | tacaactgtc | atttgatgga |
| 551 | tcaactcaaa | ataccactto | tgcttcacag | agcetteagt | gttttcttat |
| 601 | acatteres | tcatataga | cttcttcagc | aacgatctgc | aacgaaggta |
| 651 | tgattccga | ctcatagac | aaacacctgt | tgcagccatc | ccctcttccg |
| 701 | aaaccaacct | tttagagag | aaaattttct | cggggtacga | aacgctgcac |
| 751 | Gatgaatte | ctcagacgag | ctaggcattc | cagctgaaga | cgtaccagtt |
| 801 | aaaataaaa | ctcctcttgg | tcgcattctt | tacaaagctc | catctgacgg |
| 851 | accaraggga | yaycacgaac | tggactatct | tctatttatt | atcomeme |
| 901 | cyanacacya | tccaaaccca | gatgaagttg | ctgacgetaa | Otacottaat |
| 951 | cycyayyayı | cgaaagagat | actgagaaaa | gctgatgcag | Otossosoo |
| 1001 | aacaaageeg | teteettggt | ttagattggt | totogataac | **** |
| 1051 | uguggugga | ccatgtagag | gaggggaaga | ttaaggacgt | CCCCCCCCC |
| 1101 | addactatcc | acaagttgac | ttaagagaaa | qtctcttaar. | ttctactatt |
| 1151 | ragicality | ttcaataagt | ggatggtgat | gagcagtttt | tatocttcct |
| 1201 | ccaaccccgg | Cttttcaatt | tgctttatgt | gttgaacttg | taacatattt |
| | agicaaatat | gagacetege | gagttgaatt | tgaggttata | tttatagttt |
| 1251 | EGGGAACAFA. | | | | |

FIGURE 17B





Haematococcus pluvialis Ipil 1 ctcqqtaqct qqccaca

| L | ctcggtagct | ggccacaatc | GCT at t t con | | |
|------|------------|------------|----------------|--------------|------------|
| 51 | gatgccgcga | tgcttcgttc | gttgeteaga | | geggeagtee |
| 101 | cgtgaactcc | GCCCagcagc | ccagctgtgc | , adccccacac | atatecece |
| 151 | agctcaggag | | ccayctytyc | acacgcgcga | Ctccagttta |
| 201 | ctgtcgcgcg | | acgctcatgc | agcccagcat | Ctcagccaat |
| 251 | ggcaggcggg | | cacagaccac | argagggtg | caagcacctg |
| 301 | tggtggatgt | | atgagctgat | gctgaaggac | gagtgcatct |
| 351 | | | atcacaggcc | AFGCC3CC- | gctggagtgt |
| 401 | tatattatta | taccacatca | gcctgcaggc | CECCECCS | gggccttctc |
| 451 | catacasas | tttgacgatc | aggggcgact | | |
| 501 | geteadade | caccttccca | agtgtgtgga | | |
| 551 | | | | | |
| 601 | | | | | |
| 651 | | | | | |
| | | | | | |
| 701 | | | | | |
| 75.1 | | CCCGGCGC | AACCCTTT364 | 3 | |
| 801 | J-Jucyceau | ayyayctaca | GCAGATGATG. | C2 | |
| 851 | | | | | |
| 901 | | | | | |
| 951 | | | | | |
| 1001 | gacacgtcat | ggggtggaat | tocotactta | cagaagetge | aggatgtgaa |
| 1051 | tctgagactg | aacctgcagt | Caccteces | gcagcttcgt | atctcctttt |
| 1101 | gataaaatgt | | tttataaa | aaggtcaggt | aaaatggctc |
| 1151 | aaaaaaaa | aaaaa | cedede | atactgaact | ccaagaggtc |
| | | | | | |

FIGURE 18A

Haematococcus pluvialis Ipi2

| _ | | | | | |
|---|--|---|--|--|--|
| 1 51 101 151 201 251 301 351 401 451 501 551 601 751 801 851 | tggaacctgg cagaggcctc gtgcacacgc gaggaccgca gtcgcaggat acgacaacat ccacatcagc tgacgaccag ccttcccaag cagaccccag acctggcgca taccagcgca cactactgtg | cagaccacat gagctgatgc cacaggccat ctgcaggcct gggcgactgc tgtgtggacg atgaggtgga aaagctgctg ccagctgccg. ccgcggacgt gagatggact ccctgacgag agatgatgca atcgccgcgc | cgcgcgtgaa tttaagctca gaggggtgca tgaaggacga gccagcaagc gctgcaccgg tgctgcaaca aacacctgct ccaactaagc ccatccgcaa gcaagcgcgt gcagccggct acatctatt gtggacgaag gccggacaac gcttccttga | ctecgeccag ggageatgea ageacetggg gtgeatetta tggagtgeca geettetetg gegtgeaege geageeaeee caggtggeeg gttggageae ttegetteet gegacaeaat cateegggee teaggtaegt gggttgeaat gggttgeaat | caggcgggca gtggatgctg caaattccta tgttcctgtt tcaaaaatca tctacatggg acggcacagt gagctgggga cacgcgtttg cagcgctctg aacgtcacct gacgcaagag ggtcgccgtg gctgacctgg |
| 701 751 801 | gggcgagcac tggcgcccaa gagctgcggc gtttcgcatc acgcggccct atcaacgaag ggtggaattg cctgcagagc | gagatgact ccctgacgag agatgatgca | acatettatt gtggacgaag geeggacaae getteettga aaacaegagg gaagetgeag agettegtat ggtgeateat | catecgggee teaggtaegt gggttgeaat gegttggtgg attggggaac gatgtgaaga eteetttte atteategte | aacgtcacct gacgcaagag ggtcgccgtg |

FIGURE 18B



FIGURE 19A

| Lactuca s | Lactuca sativa Ipi2 | | | | | | | |
|-----------|---------------------|------------|------------|------------|------------|--|--|--|
| 1 | tattcgcttc | aaaatctctt | ccattaacto | ctcaaatctc | caccttcgcc | | | |
| 51 | ggtcttaatc | tccgccggcg | cactttcacc | accataaccg | ccaccataga | | | |
| 101 | tgacgattcc | ggcatggacg | ctgtccagag | acotctcato | tttgatgatg | | | |
| 151 | aatgcatttt | ggttgatgaa | aatgacaatg | ttcttgggca | tgataccaaa | | | |
| . 201 | tacaattgtc | acttgatgga | gaagattgag | aaagataatt | tocttcatao | | | |
| 251 | agcattcagt | gtattttat | tcaattcaaa | atacqaatta | ctccttcagc | | | |
| 301 | aaaggtcaga | aaccaaggtg | acatttcctt | tggtatggac | aaacacctgt | | | |
| 351 | tgcagccatc | cactatacag | agaatcggag | ttaattcccq | aaaatgccct | | | |
| 401 | tggggtcaga | aatgctgcac | agaggaagct | tctagatgaa | ctcggtatcc | | | |
| 451 | ctgctgaaga | tgttccagtt | gatgagttca | caactttagg | togcatgttg | | | |
| 501 | tacaaggctc | catctgatgg | aaaatggggt | gaacatgaag | ttgattacct | | | |
| 551 | actcttcctc | gtgcgtgacg | ttgccgtgaa | cccaaaccct | gatgaggtgg | | | |
| 601 | cggacattag | atacgtgaac | caagaagagt | taaaagagtt | actaaggaag | | | |
| 651 | gcggatgcgg | gtgaggaggg | tttgaaattg | tccccatggt | ttaggctagt | | | |
| 701 | ggtggacaac | ttcttgttca | aatggtggga | tcatqtccaa | aaggggacac | | | |
| 751 | tcaatgaagc | aattgacatg | aaaaccattc | ataagttgat | atgaaaaatg | | | |
| 801 | | atggtggtgg | | | | | | |
| 851 | tcqgtccttc | ttttttaac | gtttttttt | tttcttttat | tgggagtgtt | | | |
| 901 | | ttgtaacgta | | | | | | |
| 951 | | cgttaattta | | | | | | |

FIGURE 19B



Chlamydomonas reinhardtii Ipil

(Note: the isomerase cDNA probably ends at ca. base 1103; the second half of the cDNA is similar to extensin and other hydroxyproline-rich structural proteins)

| 1 | ggcacgagc | t cgagtttgt: | t ttaccatga | c atcgggaat | |
|------|------------|--------------|-------------|--------------|------------|
| 51 | aactacctca | attactcaa | taactcgcg | | |
| 101 | cgctgttttc | : tctgctccad | ctaccgage | | |
| 151 | gatgtcataa | actcccacti | | | . , |
| 201 | ccagagcgca | acctgtctta | | | gagcccaagc |
| 251 | caaagccgtg | ctctcgttg | | | gegeetegeg |
| 301 | aggactttca | caggeteaaa | | | ggccgggagc |
| 351 | cctgggaagg | • • | | | agttcgtcaa |
| 401 | tgcttggtgg | | | , acttcatgca | gegggacgag |
| 451 | cgactgccac | | | J ctaggcacco | ccaacaagta |
| 501 | accgcgcctt | | | ccaqccctgc | ggccgcctgc |
| 551 | Caccacacac | | | ccgacggccg | actgctgctg |
| 601 | Cagcagcgcg | | | ccgggtatat | ggaccaacac |
| 651 | ctgctgctcg | | cgggccagg | gccggacgag | gtggacctgc |
| 701 | cggcggcggt | | caggtgccgg | gcatcaagge | ggcggcggtg |
| 751 | cgcaagctgc | | ggggatacco | ccggagcagg | |
| 801 | ctccttctcc | 3 - | gtctgcacta | ctgcgccgcc | |
| 851 | cgcacggccc | | tggggcgagc | acgaggtgga | |
| 901 | ttcgtgcggc | cgcagcagcc | cgtcagcctg | cagcccaacc | ctacgtgctg |
| 951 | ggacgccacg | cgctacgtga | cgctgccgga | gcttcagtcc | cagacgaggt |
| | accccggcct | cagctggagc | ccctggttcc | gcatcctggc | atgatggcgg |
| 1001 | gccttcctgc | ccgcctggtg | gggcgacctg | aacccccggc | cacacagccc |
| 1051 | cggcagccga | ctgtcggact | ggggcaccat | aagcggcgct | ggcgcccggg |
| 1101 | aaaggggaag | caggggcggg | agcgggggat | ccaccgcgtc | atgtgaagaa |
| 1151 | ttgtgatgcg | gcgtgggatg | aggtctgaag | gaatgggaat | gtgaatgcga |
| 1201 | cgggcgtgag | cgtgtgtgta | cgtgagcgac | acagggggaa | aatcgggggg |
| 1251 | gcgatgggta | catgtgtgtg | cggagggtcg | aaagccggga | ggcggaccgc |
| 1301 | gcatagcgtg | ttgtgtgtgt | geggetgege | gtgggtcggt | cggttgcgcg |
| 1351 | acggaggaga | aggcacacgc | acctaccaca | gggtatgtgg | gcacccgggc |
| 1401 | ggcgggcctc | actcctggtc | aggtggcgcg | gaggtgtgtc | aggggccatg |
| 1451 | ggggctgcac | ccatatgage | gtgcccagtg | gtctcgtggg | cagagtggca |
| 1501 | tcacttggtg | aggtggggcg | ggcgcactgc | cgcgctgggc | taagtootta |
| 1551 | gaaggacacg | gtgtgtgagc | aggtggctgt | gggcggcggg | cgcagtggca |
| 1601 | ggcggatagc | | ggtggagctc | tggccgtgcc | ggccgtgagg |
| 1651 | tgcaggccgc | gatatgacgt | tgtgcttggc | cgctgtaatg | cgggagaatg |
| 1701 | cgttggggag | gagaagcggg | cggtggcagg | aggccgcagg | ctgcagcacc |
| 1751 | gggcgcctga | gtgccgcctg | caggegegge | gccgggcggg | cctgagtaat |
| 1801 | ggacgagctg | gtagtggcgg | ccacaggagg | cgcaggaggc | agcagcagga |
| 1851 | gtggccatac | gagggacccg | ttggcaaccc | | gtgtaacata |
| | 3-39ccatac | aaaaaaaa | aaaa | | |

FIGURE 20



FIGURE 21A

| Oryza . | ativa Ipil | | | | |
|--|--|--|--|--|---|
| 10 13 20 25 30 35 40 45 55 66 70 75 | agcggctcat gttgttggcc atctgaaaat aatatgaact ctagtttgga gcttatacag tcttggatga tcttggatga tcccgaaccc ctgaaggagc gtctccctgg atcacgtcga cacaagctga aagactctgt gaagtcagaa | gttcgacgac atgaatcaaa ctacttcata cctactccag ccaacacttg gaaaactacc gctgggcacc gtcggatgct cttgactacc ggacgaagtg tcatccgcaa ttccggctgg gaaaggcacc agtaaggact tcttgtgctg gaagcttttg | gcgatgttgt ctgcatatta tatgtttctg | cgggatggac tggtggatga catctgatgg tgtattcctg caacaaaggt cctctgtacc aaatgctgct atgtgccagt ccatctgatg cgtccgcgac aatacgtgag ggagaggaag cttcctcatg ccgtggacat ggctggaaag ctcttaccag | gaggtccaga acaagacaat aaaaaatcga ttcaactcaa tacatttcct gtgagtctga cagaggaagc tgaccaattc gaaaatggg gtgaaggtag ccgtgagcag gcctgaagct ggctggtggg ggagaccatc aatgatcctg ggaagttgca |
| 90 95 100 | l gcaaacttct | tgactgagag atattataca | | agagtgtcta | tgttaattta |

FIGURE 21B

| | | | *********** | |
|--|---|--|--|--|
| Plant and Green Algal Isopentenyl Pyrophosphate Isomerases (IPI) CDNAs that were isolated and identified by color complemtation in E.coli | 60 61HGDDSGHDAVQR RLHFDDECILVDECD SA AHGDSS-HDAVQR RLHFDDECILVDEND TS THGEVADAGHDAVQR RLHFDDECILVDEND TS THGEVADAGHDAVQR RLHFDDECILVDEND TS THGEVADAGHDAVQR RLHFDDECILVDEND AGA AAAVEDAGHDEVQR RLHFDDECILVDEND AAAAVEDAGHDEVQR RLHFDDECILVDEND TS THTDFRDAGHDAVQR RLHFEDECILVDEND THTDFRDAGHDAVQR RLHFEDECILVDEND TH THRASTWAG-GQSQD ELHLKDECILVDEND THRASSTWAG-GQSQD ELHLKDECILVDEND THRASSSTWAG-GQSQD ELHLKDECILVDEND | CSHPLYRES | 241 VAVNPNPDEVADIKY VGLDPNPDEVADIKY VGLDPNPDEVADIKY VKYDPNPDEVADAKY VKYOPNPDEVADAKY VKVOPNPDEVADAKY VKVOPNPDEVADAKY VKVOPNPDEVADAKY VKVOPNPDEVADKY VKVOPNPDEVADKY VKTLØPNPDEVADKY VKLØPNPDEVADKY | Tagetes erecta (marigold) Lactuca sativa (romaine lettuce) Lactuca sativa (romaine lattuce) Adonis palacatina (pheasant's eye) Adonis palacatina (pheasant's eye) Arabidopsis thaliana Atabidopsis thaliana |
| entenyl Pyropho and identified | PRECENTERS NSPRECLNELFAS NSPRECLNELFAS NSPRECLNELFAS SISANLSRAEDRIT SISANLSRAEDRIT WFGAGLSQAQSVAN | 136 SATKUTFELUWTHTC SATKUTFELUWTHTC SETKUTFELUWTHTC SATKUTFELUWTHTC SATKUTFELUWTHTC SATKUTFELUWTHTC SHTKUTFELUWTHTC SKTKUTFELUWTHTC SKTKUTFELUWTHTC ARSKITFPSUWTHTC ARSKITFPSUWTHTC | 226 EHELDYLLFIVAD EHELDYLLFIVAD EHELDYLLFIVAD EHELDYLLFIVAD EHELDYLLFIVAD EHELDYLLFIVAD EHELDYLLFIVAD EHELDYLLFIVAD EHELDYLLFIVAD EHEVDYLLFIVAD EHEWDYLLFIVAD EHEWDYLLFIVAD | 316 HKLI 232 HKLT 290 HKLT 295 HKLT 295 HKLT 295 HKLT 296 HKLT 296 |
| and Green Algal Isopentenyl Pyrophosphate that were isolated and identified by cold | 11 45 LPRKSSFPPHPS LOORCRTLSISSSIT FAHRPLSSIS LOFKLRSHOLLS LOFKLRSHOMTLHOP LAOSRALVARVSSAL | 121 VELENSKYELLLOOR VELENSKYELLLOOR VELENSKYELLLOOR VELENSKYELLLOOR VELENSKYELLLOOR VELENSKYELLLOOR VELENSKYELLLOOR VELENSKYELLLOOR VELENSKYELLLOOR VELENSKYELLLOOR VELENSKYELLLOOR VELENSKYELLLOOR VELENSKYELLLOOR | 211 225 | 301 315 316 317 317 317 317 317 317 317 317 |
| nt of Plant and ted by CONAs tha | TTKTLSASCSSPAVH TTKTLSASCSSPAVH RESLALSSSFSSFR PRVNSAQQPSCAIAR PRVNSAQQPSCAIAR PRVNSAQQPSCAIAR PRVNSAQQPSCAIAR PRVNSAQQPSCAIAR | 106 KIEFGHILHRAFS KIEKGHKLHRAFS KIEKONLLHRAFS KIEANLLHRAFS KIEASHLLHRAFS KIEASHLLHRAFS KIEASHLLHRAFS KIEASHLLHRAFS KIEASHLLHRAFS KIEASHLLHRAFS KIEASHLLHRAFS KIEQPAGLLHRAFS ANKGQPCGRLHRAFS ANKGQPCGRLHRAFS | 210 COETPLERMIY- DEFTPLERMIY- DEFTPLERMIY- DETTPLERMIY- DETTPLERMIY- DOFTPLERMIY- DOFTPLERMIY- STRETPLERMIY- STRETPLERMIY- SARRETRUHYC SARRETRUHYC SSFSFLTRUHYC | 286 286 286 286 286 286 286 286 286 286 |
| iple Sequence Alignment of I sequences were predicted by | MSSIRINDLYSIEST MSSIRINDLYSIEST MSSIRINDLYSIEST MSSIRINDLYSIEST MSSIRINGLINI | 105 NVCGHOTKYNCHLME KVCGHOTKYNCHLME KVCGHOTKYNCHLME KVCGHOSKYNCHLME KVVGHOSKYNCHLME KVVGHOSKYNCHLME KVVGHOSKYNCHLME NVCGHOSKYNCHLME NVCGHOSKYNCHLME NVCGHOSKYNCHLME NVTGHASKLECHKE NITGHASKLECHKE RLGTANKYDCHRE | 195 ORKLLDELGT PAED VRKLENEGT PAHO IRKLENEGT PAHO IRKL | PECELKISPHERIN PECHKISPHERIN ECCIKISPHERIV E |
| Clustalw 1.7 Multiple These amino acid seque | T.erecta 1 L.sativa 1 L.sativa 2 A.palaestina 2 A.palaestina 1 A.thaliana 1 A.thaliana 2 H.pluvialis 1 H.pluvialis 2 C.reinhardtil 1 | T.erecta 1 L.sativa 1 L.sativa 2 A.palaestina 2 A.palaestina 1 O.sativa 1 A.thaliana 1 A.thaliana 2 H.pluvialis 1 C.ceinhardtii 1 | T.erecta 1 L.sativa 1 L.sativa 2 L.sativa 2 A.palaestina 2 O.sativa 1 A.thaliana 1 A.thaliana 3 H.pluvialis 2 R.pluvialis 2 C.reinhardtil 1 | T. erecta 1 6 6 1 1 2 attiva 1 6 6 1 1 2 attiva 1 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 |
| These | 110087664 | | | |

```
Comparison using GAP program of the Genetics Computer Group
         Gap Weight:
                             50
                                      Average Match: 10.000
        Length Weight:
                                      Average Mismatch:
                                                              0.000
               Quality:
                            17392
                                                   Length:
                                                               1904
                 Ratio:
                            9.411
                                                     Gaps:
Percent Similarity: 95.331
                                    Percent Identity: 95.331
Match display thresholds for the alignment(s): | = IDENTITY
 Adonis palaestina E-cyclase #3 x Adonis palaestina E-cyclase #5
        1 gagagaaaaigagtgttatittaatgttactgtcgcattcttgcaacac: 49
1 .....aaaggagtgttctattaatgttactgtcgcattcttgcaacact 44
      99 tga.cggagiatctagctaiggaactactiggtgttcgcaacctcat
95 tgagcagagtatctggctatggaactacttggtgttcgcaacctcat
                tgccctgtctggacttttggaacaagaaaccttagtagttcaaaac 197
tgccctgtgtggacttttggaacaagaaaccttagtagttcaaaac 194
      148
      145
     248 gtgagggctgatggtggaagcgggagtagaacttctgttgcttataaaga 297
245 gtgagagctgatggtggaagcgggagtagaagttctgttgcttataaaga 294
     298 gggttttgtggatgaagaggattttatcaaagctggtggttctgagcttt
          tgtttgtccaaatgcagcaaacaaagtctatggagaaacaggccaagctc
     348
     345
          gccgataagttgccaccaataccttttggagaatccgtgatggacttggt 447
     398
     395
          tgtaataggttgtggacctgctggtctttcactggctgcagaagctgcta 497
     448
     445
     498 agctagggttgaaagttggccttattggtcctgatcttccttttacaaat
    548 aattatggtgtgtgggaagacgagttcaaagatcttggacttgaacgttg
          tatcqaqcatqcttqqaaqqacaccatcqtatatcttqacaatqatqctctatccgaqcatqcttggaaqqacaccatcqtatatcttgataatqatqctc
     598
     595
    648 ctgtccttattggtcgtgcatatggacgagttagccggcatttgctacat 697 645 ctgtccttattggtcgtgcatatggacgagttagtcgacatttgctacat 694
    698 qaaqaqttqctqaaaaqqtqtqtcqaqtcaqqtqtatcatatctqaattc
         taaagtggaaaggatcactgaagctggtgatggccatagccttgtagt
    798 grgaaaacgacarctttatcccttgcaggcttgctactgttgcatctgga 847
795 grgaaaatgagatctttatcccttgcaggcttgctactgttgcatctgga 844
```

FIGURE 23

VO 99/63055

FIGURE 23 (cont.)

| 84 | 8 gg | age | 775 | çaç | Jgg | aaa | ąçţ | Çţ | t q | ga | ata | ato | aaa | iat | ac | ia t | aa. | -:- | | | | | |
|--------------|-------------|---------------------------------------|---|-------|---------------------------------------|-----------------|-----------|--|---|--|---|-------------|---|------------|---|--|-------|--|-------------------------------------|---|----------------------|--|--------------|
| 84 | 5 ჭბ | :aģd | FFF | cad | 399 | aaa | lct | ff | få | ga | ďť. | | ITI | | | ֓֓֓֓֓֓֓֓֓֓֓֟֝֟֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓ | 77 | ŢŢŢ | ֓֞֞֞֞֞֓֓֓֞֜֝֓֓֓֓֞֜֟֜֓֡֡֡֓֡֓֡֡֡֡֡֡֡֡ | 11 | | 98 | 97 |
| 89 | 8 tg | 1799 | aa | ąçi | gç | ţţ; | Fg | gt | ġţ | gg | ago | a t | gá | ac | ito | iga | ga: | ica | a F | ge | itac itac | g B | 94 |
| 89 | 5 tġ | tçç | àà | åċc | gċ | ff | ltģ | ģg | ġέ | gg. | agg | 15 | r l | ומס | ίξ | ΙďΑ | ĮΪ. | ֓֞֞֞֞֓֞֓֓֓֞֟֓֓֓֓֓֟֓֓֓֟֡֓֓֓֓֡֓֡֓֡֡֡֡֡֓֡֓֡֡֡֡֡֡֓֡֓֡֡֡֡֡֓֡֓֡֡֡֡֡֡ | וַנְ | ֓֞֟֝֓֓֟֝֟֝֟֟֝֟֝֟֟֝֟֟֝֟֟֝֟֝֟֟֝֟֟֝֟֝֟֝֟֝֟֝ | | 9 | 47 |
| 941 | 8 at | 999 | aa. | fff | aa | t gg | JF A | FF | ŧа | ţg: | gaç | - - - | çå | ga | ga | ct | ata | ita | ca | aci | 2022 | 3 9 | 44 97 |
| 945 | 5 át | ĊĊĊ | aa | ctt | àà | tģģ | ıta | ff | cå | Łģι | gad | ta | da | ga | ga | ęξ | I [] | ίξ | IJ | ֓֞֞֓֞֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֡֓֓֓֓֡֓֡֓֓֡֓֡֓֡֓֡֓֡֓֡֓ | gaa. | ֓֞֞֞֜֞֓֓֓֓֞֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֡֓֓֡֓֡֓֓֓֡֓֡֓֡֓֡ | 9 / 9 4 |
| 998 | | aca | 950 | 766 | çg | gaa | ga | ag | ġą | ţą | FFF | :aa | çạ | ţţ | FÇ | ţÇ | tat | ġt | ca | tac | igaa. CCa CCa | . J. | 94 047 |
| 995 | tt | àċā | ġt | ġċt | ċġ | ģáá | gá | άģ | åå | ta | FÇÇ | :dd | ca | ff | fc | fc | Łał | äξ | ĮΙ. | | III | 10 |) 4 4 |
| 1048 | . <u>11</u> | 599 | Ça | rca | aga | PFF | FF | 77 | ۴F | tga | 799 | raa | ąċ | ۲ŧ | gţ | FF | 399 | Ċţ | ça | aaa | gato | 10 | 97 |
| 1098 | gt | cgc | caa | aca | aga | act | .tti | tt | tt. | tġa | àġġ | ää | άċ | ċŧ | ąŧ | ff | βġċ | çŧ | ζå | aaa | gate | 10 | 94 |
| 1095 | ĬĬ | ֓֓֟֓֓֓֟֓֓֓֓֟֓֓֓֓֓֟֓֓֓֟֓֓֓֓֟֓֓֓֓֓֓֓֓֓֓ | | ĪĪ | Ç | ΪŢ | II | ֓֓֞֜֞֜֓֓֓֓֓֓֓֜֡֓֓֡֓֓֡֓֡֓֡֓֜֜֡֓֡֓֡֡֡֡֡֓֡֓֡֓֡֡֡֡֡֡ | ga | aga | ga | aa | ą¢ | ţa | ąţ | gta | FAC | ġą | FF | jąą | gato gact gact | 11 | 47 |
| 1148 | CE | aca | + = + | | cga | 110 | tac | Σ τ (| gaa | aga | aga | àà | àĊ. | tg | àt | ġŧ | tac | ģå | έŁς | Jaa | gact | 11 | 44 |
| 1145 | II | קקק ממם | ֡֓֓֓֟֟֞֟֓֓֓֟֟֓֓֓֓֓֟֟֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓ | | ֓֓֓֟֓֓֟֓֓֓֟֓֓֓֟֓֓֓֟֓֓֓֟֓֓֓֓֓֓֓֓֓֓֓֓֓֓ | [[] | | ֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓ | ֓֟֟֓֓֟֟֓֓֟֟֓֓֟֓֓֓֟֓֓֓֟֓֓֓֟֓֓֓֟֓֓֓֓֓֓֟֓֓֓֓ | | Ī | 19 | aa. | ga | 99 | aat | 99 | ŧ۴ | 2 | ta | gace FFC EEC | 11 | .97 |
| 1198 | ta | tta | aac | at | ter | , | acc | - 2 | a aq | 300 | :ta | tg | aa -: | ga | ġġ. | äät | ġġ | tċ. | ata | ta | ffç | 11 | .94 |
| 1195 | L Eg | ffå | gto | | ſΪ | ΞĮ | | | ֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓ | ֓֓֓֓֟֟֓֓֟֟֓֓֟֟֓֓֓֟֟֓֓֟֝֟֓֓֟֓֓֟֝ <u>֚֚֚֟֝</u> | ֝֟֓֓֟֟֓֓֟֟֓֓֓֟֟֓֓֓֟֟֓֓֓֓֟֝ <u>֚֚</u> | 17 | ֓֟֝֟֓֓֟֟֓֟֓֓֓֟֟֓֓֓֓֟֓֓֟֓֓֓֟֓֓֓֓֓֟֓֓֓֟֓֓ | | ag: | aac | 9 | ago | rat | ++ | ttėć ggtg ggtg | 12 | 47 |
| 1248 | Çţ | gça | gça | aġ | çat | aa | tac | at | ico | ac | .ca | ay. | aàd | 3 4 | aga | aac | ict | ago | cat | tt | gġtġ | 12 | 44 |
| 1245 | çt | jċa: | gca | dg | <u>La l</u> | .gg | | a | | ao | ΙŢ | Ĭľ. | | | ֓֓֓֓֓֟֓֓֓֓֟֟֓֓֓֓֓֓֟֟֓֓֓֟֟֓֓֓֓֓֓֓֓֓֟֓֓֓֟֓֓֓֓ | בנו | 79 | | IF | içg | ggtg atçå gtca | 12 | 97 |
| 1298 | Sta | a t ç | aga | ąġ | çţç | ça | aāa | ţ | itç | Çţ | tç | ta | tåa | 3 C 1 | ta: | caa | ag. | i L | - - - | cg | gtca agca agca | | |
| 1295 | çtç | gt ċ. | àġå | aģ | ffc | :ca | aaa | £ | fξ | ίς f | ξĻ | ξģι | Lad | 15 | | aa | II. | ֓֓֓֓֟֓֓֓֓֓֓֟֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓ | | ֓֟֓֓֓֓֟֟֓֓֓֟֟֓֓֓֓֓֟֓֓֓֟֓֓֓֟֓֓֓֟֓֓֓֟֓֓֓ | | 13 13 | |
| 1348 | aga | ta. | PFF | cf | a ça | ta | fgf | gç | 7 5 5 | FF | F g | gad | ;åą | aç | JC | gt | ac. | iat | aa: | ac | ayca arri | 13 | _ |
| 1345 | aga | taa | ict | ċŧċ | ģċg | tai | tġt | ģģ | 144 | ŧċ | fģ | ģåd | al | lac | jt | ďξ | ge | i [| II | II. | III | 13 | |
| 1398 1395 | Çaş | IEgo | aa T | 9¢9 | 759 | gag | 7Ç4 | 90 | 95 | FF | 999 | 799 | aåa | gç | jąç | çq | aaa | jąç | gţ | çaa | attt aga aga | 14 | |
| 1448 | caa | itgo | caa | gca | atġ | gaç | jċa | ġt | Ċt | tt | ġġ | -c | ààà | ιģģ | jåģ | ι¢ģ | aaa | lac | ġŁ | caa | laga | 14 | |
| 1445 | Y | | | M | 11 | Feg | Igg | Ħ | 49 | 99 | 951 | FA | ۲ç | 159 | rça | 96 | tac | jąç | ąţ | t ga | agç | 14 | 97 |
| 1498 | aac | Cac | 122 | cåt | . t t | • • • | gga | -: | ag | ag | cti | àt | :tġ | įtģ | ŗċá | ġċ | tad | ját | åŧ | Łġź | aga agç agç | 149 | 94 |
| 1495 | III | Cac | TA | | ŢŢ | וון | ַקָּאָ | ŢŢ | Ï | II | [[] | | 79 | | 99 | ça | act | 79 | ga | F9 1 | agc ggt | 154 | |
| 1548 | ggg | gtt | tc | ctt | aa | ato | tt: | cá | ct | 200 | cat | | :gc | ככ | gc | ca | act | tġ | ġà | tġt | ġġt. | 154 | |
| 1545 | 999 | åff | ξÇ | fff | gg | ΪĘ | ΞĘ | IJ | ΪĪ | ĬĮ. | ֓֞֟֞֟֓֓֟֟֓֟֓֟֟֓֟֓֟֓֟֓֟֟֓֓֓֟֓֓֓֟֓֓֓֟֓֓֓֟ | [נֻ] | 11 | | Ţ | II | | a | F9 | 777 | FFF | 159 | |
| 1598 | arg | ţąç | aţ | ą į t | tg. | FFF | ţg | σċ | CÇ | Ca | aac | ao | ċa | ta | ac | ma: | tai | .C. | cg c+ | | :CCC | 159 | |
| 1595 | åtģ | táć | at | åff | ŧφ. | fff | :£ģ | ĝς | gċ | ca. | aad | :do | ca | | I | ď | | | II. | III | gaġ gaġ | 164 164 | |
| 1648 | aca aca | FFF | gçi | FFF | ça | gat | ÇÇ | ţţ | ÇF | ggi | t g ç | :ag | Ėŧ | aţ | aa | tt | aaå | ac | tt: | acc | tcá | 169 | |
| 1645 | à¢á | ttt | ġċſ | ttt | Ċå | ģát | င်ငံ | ŧŧ | çŁ | ģģί | Fåç | :dg | ff | ۵Ł | άđ | tai | aga | 4C | fξ | | | 169 | |
| 1698 | 444 | 995 | aaı | řŶ٠ | 1 | fgf | ידדי | ţġ | ţg | aaa | 795 | aç | ġg | ¢9 | ţc | ţç | att | ąą. | ata | aaa | tgå Ega | 174 | - |
| 1695 | aaa | ggt | àgt | tet | cai | tct | ati | tá | ŧt. | ààa | ict | ct | άģ | ŧģ | ŧτ | fç | acc | aa. | a Ła | lala | Ega | 174 | |
| 1745 1745 | gga gga | ĬĨ | FF | 95 | ata [] | 159 | Fai | FÃ | 59 | ato | PPF | çţ | ¢ţ. | ąţ | 95 | 459 | ţċ | Çţ. | aţa | a t ţ | çţġ | 179 | |
| 1795 | yyd atc | tc: | tar | -ga | ato | gtg | | ca | tg. | ato | at | ĊĖ | ċŧ. | at | ġŧ | àta | tc | çŧ | gŁ | jcę | çfå | 179 | 4 |
| 1795 | III | ዮዮዋ | ትዋና | 44 | 649 | 4 C C | gaa | aa | atı | ÇCa | 3 C C | ga | ta | ga. | aa. | aaa | خده | aa. | 223 | | a | 184 | |
| 1845 | aaa | a. | | | | 14 C | 900 | -g | 33 1 | בכנ | .ga | ta | t t | gt | tg | tgt | cà | áå | CC | 3 90 | caa | 184 | |
| 1845 | tga | tat | aaa | ıat | aaa | | tat | · · | 729 | • • • | • • • | • • | •• | • • | •• | | • • • | • • | • • • | • • • | • • • | 184 | _ |
| | , | | | 7 | | | | | 901 | -40 | .ad | a a | g۲ | ag: | CC | ててり | :ct | CCI | Cta | aaa | aaa | 189 | 4 |

WO 99/63055

GAP program of Genetics Computer Group blosum62.cmp Gap Weight: 12 Average Match: 2.912 Length Weight: 4 Average Mismatch: -2.003 Quality: 2728 Length: 530 5.147 Ratio: Gaps: 0 Percent Similarity: 99.623 Percent Identity: 99.057 Match display thresholds for the alignment(s): | = IDENTITY Adonis palaestina E-cyclase #3 x Adonis palaestina E-cyclase #5 1 MELLGVRNLISSCPVWTFGTRNLSSSKLAYNIHRYGSSCRVDFQVRADGG 50 1 MELLGVRNLISSCPVWTFGTRNLSSSKLAYNIHRYGSSCRVDFQVRADGG 50 51 SGSRTSVAYKEGFVDEEDFIKAGGSELLFVQMQQTKSMEKQAKLADKLPP 100 51 SGSRSSVAYKEGFVDEEDFIKAGGSELLFVQMQQTKSMEKQAKLADKLPP 100 101 IPFGESVMDLVVIGCGPAGLSLAAEAAKLGLKVGLIGPDLPFTNNYGVWE 150 101 IPFGESVMDLVVIGCGPAGLSLAAEAAKLGLKVGLIGPDLPFTNNYGVWE 150 151 DEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKR 200 151 DEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKR 200 201 CVESGVSYLNSKVERITEAGDGHSLVVCENDIFIPCRLATVASGAASGKL 250 201 CVESGVSYLDSKVERITEAGDGHSLVVCENEIFIPCRLATVASGAASGKL 250 251 LEYEVGGPRVCVQTAYGVEVEVENNPYDPNLMVFMDYRDYMQQKLQCSEE 300 251 LEYEVGGPRVCVQTAYGVEVEVENNPYDPNLMVFMDYRDYMQQKLQCSEE 300 301 EYPTFLYVMPMSPTRLFFEETCLASKDAMPFDLLKRKLMSRLKTLGIQVT 350 301 EYPTFLYVMPMSPTRLFFEETCLASKDAMPFDLLKRKLMSRLKTLGIQVT 350 351 KIYEEEWSYIPVGGSLPNTEQKNLAFGAAASMVHPATGYSVVRSLSEAPK 400 351 KVYEEEWSYIPVGGSLPNTEQKNLAFGAAASMVHPATGYSVVRSLSEAPK 400 401 YASVIAKILKQDNSAYVVSGQSSAVNISMQAWSSLWPKERKRQRAFFLFG 450 401 YASVIAKILKQDNSAYVVSGQSSAVNISMQAWSSLWPKERKRQRAFFLFG 450 451 LELIVQLDIEATRTFFRTFFRLPTWMWWGFLGSSLSSFDLVLFSMYMFVL 500 451 LELIVQLDIEATRTFFRTFFRLPTWMWWGFLGSSLSSFDLVLFSMYMFVL 500 501 APNSMRMSLVRHLLSDPSGAVMVKAYLER* 530 1111111111111111111111111111111111111 501 APNSMRMSLVRHLLSDPSGAVMVRAYLER* 530

| ر د |) | |
|-----------|---|--|
| | | |
| 7 1 2 1 2 | 5 | |
| Ē | 1 | |
| | | |
| | | |
| | | |

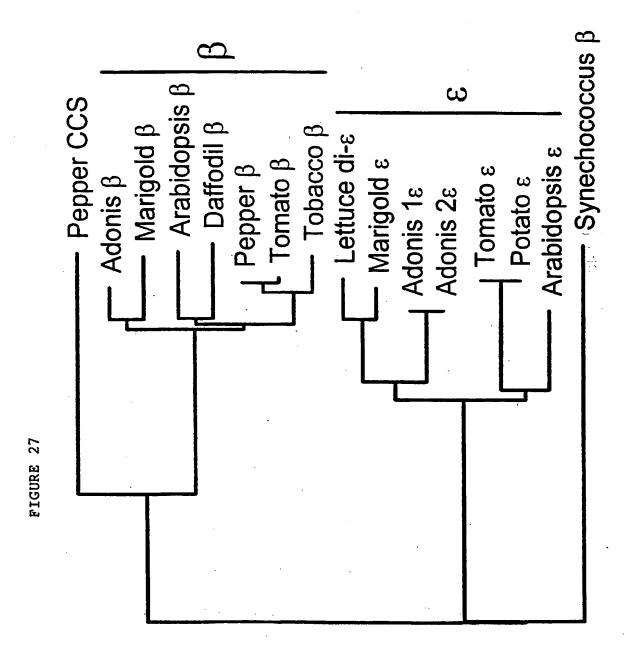
| | 103 102 102 102 100 84 72 71 71 71 | 205 205 206 206 206 206 1181 1186 1186 1186 1186 1186 1186 11 | 161 312 311 311 311 316 316 289 289 289 289 |
|---------|--|--|--|
| | 1 80 0, 0, 0, 10 0, 12 13 13 13 13 13 13 | | 1101 1111 1111 1111 1111 1111 1111 111 |
| | The state of the s | 22 VERY COURS OF COU | |
| | GGGSSGSESCVAVREDF - ADBEDFYKAGGBEILFYÖHÖNKDHDEĞSKEVDKU ADGGSGSRSSVAYKEGF - VDBEDFIKAGGBEILFYÖHÖOTKSHEKÇAKİBAKL ADGGSGSRSSVAYKEGF - VDBEDFIKAGGBELLFYÖHÖOTKSHEKÇAKİBAKL CASAKSDR - CVVDKQGISVADBEDFYKAGGBELFYÖHÖOTKSHEKÇAKİBAKL COSAKSDR - CVVDKQGISVADBEDYYKAGGBELFYÖHÖOTKSHESÜSKL CONSSGGSBCVVDKEDF - ADBEDYYKAGGBELFYÖHÖOKKOHDOĞSKISEKL - RUCSNNPYHSRVRLQVKKRAİKIV SSVVSGSAALLDLVPETKKEN - AVKVSAFSSVKSQKFGAKKFCEGLG SRSVCVKGSSSALLELVPETKKEN - AVKASTFRSEKHHNFGSRKFCETL GRSVCVKGSSSALLELVPETKKEN - AVKASTFRSEKHHNFGSRKFCETL GRSVCVKGSSSALLELVPETKKEN NINNINQLNQSKSQPOPFRFGKKSQFKLGQKYCVKASSSALELUVPETKKEN - KRHFLSPSPNPQNPNFKFFSKKPYQKKCRNGYIGVSSNQLLDLVPETKKEH - KRHFLSPSPNPQNPNFKFFSKKPYQKKCRNGYIGVSSNQLLDLVPETKKEH | HEELLKEOVE HEELLKOVE HEELLKOVE HEELLKOVE HEELLKOVE HEELLKOVE HEELLKOVE HEELLKOVE KSKAMOKOLIN KSKAMOKOLIN KSKAMOKOLIN KSKAMOKOLIN KSKAMOKOLIN KSKAMOKOLIN KSKAMOKOLIN KSKAMOKOLIN | PTELY PTELY PTELY PTELY PTELY PTELY PTELY PTELY PTELY PTELY PTELY |
| 100 | TOTO TO THE TOTO THE TOTO THE TOTO THE TOTO THE TOTO THE TOTO THE TOTO THE TOTO THE TOTO THE TOTO THE TOTO THE TOTO THE TOTO THE TOTO THE TOTO THE TOTO THE TOTO THE TOTO THE THE TOTO THE THE TOTO THE THE TOTO THE THE TOTO THE THE TOTO THE THE TOTO THE THE TOTO THE THE TOTO THE THE TOTO THE THE TOTO THE THE TOTO THE T | HITTER XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX | <u> </u> |
| | SAND SAND SAND SAND SAND SAND SAND SAND | SERLI SERLI SERLI SERLI SERLI SERLI SERLI SERCI SE SERCI SERCI SERCI SERCI SE SERCI SE SERCI SE SERCI SE SERCI SE SERCI SE SERCI SE SERCI SE SERCI SE SERCI SERCI SERCI | SLEANSERNS ERNNS E |
| | GOOT GOOT GOOT GOOT SSAL SSAL SSAL SSAL SSAL SSAL | | 320 Year Strain |
| • | TOO TOO TOO TOO TOO TOO TOO TOO TOO TOO | | VR TN MQ MQ SK ER |
| | KAGGBELL KAGGBELL KAGGBELL KAGGBELL SSS PCRNGIN SSS SSS SSS SSS SSS SSS SSS SSS SSS SSS SSS SSS ST | DDDEILIE DDDEELLE DDDDEELLE DDDDEELLE DDDDEELLE DDDDEELLE DDDDEELLE DDDDEELLE DDDDEELLE DDDDEELLE DDDDEELLE DDDDEELLE DDDDEELLE DDDDDDEELLE DDDDEELLE DDDDEELLE DDDDEELLE DDDDEELLE DDDDEELLE DDDDDEELLE DDDDDEELLE DDDDDEELLE DDDDDEELLE DDDDDDDD | |
| 80 | KAGG KAGG KAGG KAGG KAGG C C C C C C C C C C C C C C C C C C | | |
| | - ADĞEĞEYEĞ - VDĞEĞETİĞ - VDĞEĞETİĞ - ADĞEĞEYİĞ - ADĞEĞEYİĞ - ABĞEĞEYİĞ - KRAİKİV- KRAİKİV- KRÇEĞIĞ- SKKCEĞIĞ- SKÇCEĞIĞ- SKÇCEĞIĞ- SKÇCENIĞ- KIÇENIĞE | | 300 SLLWVE NLWVE SLLWVE SLLWVE SLLWVE SKWVE SKWVE SKWVE |
| | ADEEDFUVDEEDFUVDEEDFU- SVADEEDFUADEEDFUBEEDFUREEDFU | WRDII WRDII WRDII WRDII WRDII WRDII WRDII WSGAA WSGAA | |
| • | ADGGSSGSRSCVAVREDF ADEEDFVR ADGGSGSRSSVAYKEGF VDEEDFIK ADGGSGSRTSVAYKEGF VDEEDFIK CCSAKSDR CVVDKGGISVADEEDFYK CCNSSGGSDSCVVDKEDF ADEEDYYK CCNSSGGSDSCVVDKEDF ADEEDYYK RLCSNNDYHSRVRLGVKKRAIKIV KRQHLVSTSKLQNQVFRIASRNIH- VKASSFRSVKRHNFGSRKFCETL- VKASSFRSVKPHKFGSRKFCKTL- NLNNLNQLNQSKSQFQPKRFGPKKSQF | 180 CIEHO CIEHO CIEHO CIETO CIDTO CIDAN CIDAN CIDAN CIDAN CIDAN | WENSE WENS WENS |
| | 7AVRI 7AYKI 7AYKI 7AYKI 7VDKI 7VVKI 1SKUP | | |
| 09 | SESC SESC SESC SESC SESC SESC SESC SESC | DEFROIG DEFROIG DEFROIG DEFROIG DEFROIG DEFRAND DEFRAND DEFRAND DEFRAND DEFRAND | AYGVE AYGVE AYGVE AYGILA AYGILA AYGILA AYGILA AYGILA AYGILA AYGILA |
| | SSSGS SGS SGSGS SGSGS SGSGS SGSGS SGSGS SGSGS SGSGS SGSGS SGSGS SGSGS SGS SGSGS SGSGS SGSGS SGSGS SGSGS SGSGS SGSGS SGSGS SGSGS SGSGS SGS SGSGS SGS SGS SGSGS S SGS | OBE OBE OBE OBE OBE OBE OBE OBE OBE | |
| | CSVRASGGGSSGSRSCVANTEDFADEEDFYKAGGBILLE CRVDFQVRADGGSSRSSVANTEGFVDEEDFIKAGGBILLE CRVDFQVRADGGSSRTSVANTEGFVDEEDFIKAGGBILLE SSSLRQINCSAKSDRCVVDKQGISVADEEDYVKAGGBILFE RQYESIKCHSSGSDSCVVDKEDFADEEDYVKAGGBILFE IPQFHGFBRLCSNNPYHSRVRLGVKRRAIKIVSSV PTLHGFAEKQHLVSTSKLQNQVFRIASRNIHPCRNGTV HGFGAVKASTFRSKHHNFGSRKFCEGLGSRSVCV NPHHGPAVKASTFRSKHHNFGSRKFCEGLGGRSVCV HPWHGFSKQULVSTSKLQNGVFRIASRNIHPCRNGTV HPWHGFSAVKASTFRSKHHNFGSRKFCEGLGGRSVCV HPWHGFSKRHFLSPSPNPQNPNFKFFGRKGGKYCV | O MINYGUM MINYGUM MINYGUM MINYGUM MINYGUM MINYGUM MINYGUM MINYGUM MINYGUM MINYGUM MINYGUM MINYGUM MINYGUM MINYGUM MINYGUM MINYGUM MINYGUM | |
| • | SSV-RAS RVDPQVP RVDFQVP SSSLRQIK QV-ESS PQFHGPB TLHGPAHGFG PHHGP PVHGFS | | |
| | SCRV SCRV SCRV SCRV LPTL LPTL LPTL LHPY UHPS | ALIND KLIND KLIND KLIND KLIND KLIND KLIND | |
| 40 | RRFPVVRRYSTRNIRFGL-CSVRASGGGSSGSRSCVAVREDFADĒEĎFVKĀĞGBEILFVÖHĞOKRDHDBÖSKÍVDKI TRNISSSKLAYNIHRYGSSCRVDFQVRADGGSGSRSSVAYKEGFVDĒEĎFIKĀĞĞBILLFVÖHĞOKKHENDKIADKI TRNISSSKLAYNIHRYGSSCRVDFQVRADGGSGSRTSVAYKEGFVDĒEĎFIKĀĞĞBILLFVÖHĞOTKSHEKÖAKIADKI TRHKFBLLKQRRFTNLSA-SSSLRQIKCSAKSDRCVVDKQGISVADĒEĎTVKĀĞĞBILFVÖHĞOTKSHEKÖK FWSGGELCQEKSIFLAY-EQYESKCNSSSGSDSCVVDKEDFADĒEĎTVKĀĞĞBILFVÖHĞOKKOHDOÖSKISEKI HDTLLKTPNKLDFFIPQFHGPERLCSNNPYHSRVRLGVKRAIKIVSSVVSGSAALLDLVPETKKEN HDTLLKTPNKLELPTLHGPAEKQHLVSTSKLONOVFRIASRNIHPCRNGTVKAGSSALLELVPETKKEN HDTLLKTPNKLEFLNPHGPAVVSAFSSVRSOKFGAKKFCEGLGSRSVCVKASSSALLELVPETKKEN HDTLLKTPNKLEFLHPVHGPSAVVASTFRSKHHNFGSRKFCETLGRSVCVKASSSALLELVPETKKEN HDTLLKTPNKLEFLHPVHGPSAVVASSFNSVRPHKFGSRKICENWGKGVCVKASSSALLELVPETKKEN HDTLLKTPNKLEFLHPVHGPSKASSFNSVRPHKFGSRKICENWGKGVCVKASSSALLELVPETKKEN HDTLLKTPNKLEFLHPVHGPSKASSFNSVRPHKFGSRKICENWGKGVCVKASSSALLELVPETKKEN HDTLLKTPNKLEFLHPVHGPSKRASSFNSVRPKFGFKRSOFKLGQKYCVKASSSALLELVPETKKEN | | |
| | YERNI YERNI YERFI YERFI YETPN KIPN KIPN KIPN | CSILIVATION OF STREET | |
| | SKLA SKLA SKLA SKLA CLEKQ DTEL DTEL DTEL DTEL | | PECTAL PECTAL PECTAL PECTAL PECTAL PECTAL PECTAL PATAVLE PATAV |
| • | MA W W W W W W W W | 140 AAESKE AAESK | |
| | | LIANO VAOO CETS | WECEGEDON WECEGEDIN |
| 20 | SWS-VWT-VWT-VWT-RFTD | * CGBAGLAIL CGBA | TACEBER TACEBER TECEBE |
| 7 | HECVGARNP-AAMAVSTPESWB-CF HELLGVRNLISSCRWT-PG HELLGVRNLISSCRWT-PG HECVGVQNV-GAMAVITEPRIDGY HECVGVQNV-GAMAVITEPRIN HSWRAG-HHTATHAAFICERN | | GVLTINSKUDRIVERTRÖHSIDEGEDOVVI GVSTISKUDSITERSDIRITARGEDINVI GVSTIDSKURRITERSDIRITARGEDINVI GVSTISKURRITERSDIRITARGESIVVERDITE GVSTISKURRITER PROGESIVE GEGOVVI GVSTISKURRITER PROGESIVE GEGOVVI GVSTISKURRITER PROGESIVE GEGOVVI GVSTISKURRITER PROGESIVE GEGOVVI GVKFHQARVIRVIH - BESKSHLIGNDGITH GVKFHQARVIRVIH - BESKSHLIGNDGITH GVKFHQARVIRVIH - BESKSHLIGNDGITH GVKFHQARVIRVIH - BESKSHLIGNDGITH GVKFHQARVIRVIH - BESKSHLIGNDGITH GVKFHQARVIRVIH - BESKSHLIGNDGITH GVKFHQARVIRVIH - BESKSHLIGNDGITH |
| | ANAV THAV ANAV THAA | 120 IGBSVPDLVVIG FGBSVPDLVVIG IGBSVPDLVVIG IGGGDSNCILDLVVIG AGQTVLDLVVIG PLYDTSKSQVVDLALVV PAYDPSKGVVVDLAVV PHYDPSKGVVVDLAVV PHYDPSKGVVVDLAVV PHYDPSKGVVVDLAVV PHYDPSKGVVVDLAVV PHYDPSKGVVVDLAVV PHYDPSKGVVV | |
| + ; | NIL NIL NIL NIL NIL NIL NIL NIL NIL NIL | 120 IGBOALDHVV FGESVAGLVV IGOTVLGLVV AGQTVLGLVV AGQTVLGLVV AGQTVLGLVV PLYDPSKGVVVDLAY PHYDPSKGVVVDLAV PHYDPSKGVVVDLAV PHYDPSKGLVVDLAV PHYDPSKGLVVDLAV PHYDPSKGLVVDLAV PHYDPSKGLVVDLAV PHYDPSKGLVVDLAV | AREA THE TENTE T |
| ; | VGAR LGVR FGAR VGVQ | B B B GDSN GDSN GDSN PSKG PSKG PSKG | Thiskyor Therese Therese Therese Therese Therese Therese Therese Therese Therese Therese Therese Therese Therese Therese |
| ; | | FG IGG IGG IGG IGG IGGG | |
| • | | 00 00 00 00 00 00 00 00 00 00 00 00 | |
| 50 | Arabidopsiss Adonissi Adonissi Lettucess Tomatos Arabidopsiss Adoniss Peppers Tomatos Tobaccos Marigolds | Potatos Arabidopsiss Adoniss1 Adoniss2 Lettucess Tomatos Arabidopsiss Adoniss Tobaccos Marigolds | Potatos Arabidopsiss Adonissi Adonissi Comatos Arabidopsiss Adoniss Peppers Poppers Comatos Comatos |
| PotatoB | Arabidops AdonisEl AdonisEZ LettuceEE TomatoB Arabidops AdonisB PepperB TomatoB TobaccoB MarigoldB | PotatoB Arabidops AdonisB1 AdonisB2 LettuceBB TomatoB Arabidops Arabidops Arabidops Arabidops Arabidops Arabidops Arabidops Arabidops Arabidops Arabidops Arabidops Arabidops Arabidops Arabidops Arabidops DeffodilB | Potatos Arabidopsiss Adonissi Adonissi Tomatos Arabidopsiss Arabidopsiss Adoniss Peppers Tomatos Tobaccos |
| Po | Add Add Add Add Add Add Add Add Add Add | Pot Ado Ado Let Tom Adoi Pepi Tobe Hard Daff | Potatos Arabidopi Adonissi Adonissi Lettucessi Lettucessi Tomatosi Adonissi Peppersi Tomatosi Tobaccosi |
| | | | |

| $\overline{}$ |
|---------------|
| |
| u |
| cont |
| ≍ |
| 2 |
| U |
| $\overline{}$ |
| |
| 25 |
| N |
| |
| ы |
| |
| 座 |
| \Box |
| G |
| IGUR |
| بع |
| |
| |

| | 27 / 30 | # ¹ 2. |
|----------|--|--|
| | 2 | 8488888614880048 |
| | 2 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 | 378 5228 5229 5239 5239 5200 5000 5000 50110 |
| 9 | % : 0 0 0 1 0 % : ! ! ! A ! | *!!!!!!! |
| 4 | RONHSKNMLTSS REETTKOINS- KODNSATVVSGG KODNSATVVSGG RODGSKRMISLG ROHYSKNMLTS- GRUNSKQMLDHG GSPSSNS GSPSSNS GSPSSN GSPSSN GSPSSN GSPSSN GSPSSN GSPSSN GSPSSN GSPSSN GSPSSN GSPSSN | KLTF- CLOOK CLOOK CLOOK CLOOK CLOOK CLOOK CLOOK CLOOK CLOOK CLOOK |
| | KNY KOJ KRIA KRIA KRIA KRIA KRIA KRIA KRIA KRIA | 40 ATLIRTKLTF ATMIKTKLKV AVMVRAKLER AVMVRAKLER ATMIKAKLT ATLIRTKLT ATLIRTKLT ATMINALUD VWMINNLUD VWMINNLUD VWMINNLUD VWMINNLUD VWMINNLUD VWMINNLUD VWMINNLUD VWMINNLUD VWMINNLUD |
| | RONHSK REETTK KODNSA RODNSA RODNSK GSPSSN GSPSSN GSERSH GSERSH GSERSH | ATLIRT ATMIKT AVMVKA ATMVKA ATLIRT GTMLKA AKMINN VNMENNI VNMINNI VNMINNI VNMINNI VNMINNI VNMINNI |
| ٠. | | FTGATLIRTKLTF PTGATMIKTKLKV PSGAVMVRAKLER PTGATMVKAKLTI PTGATLIRTKLT PTGGTMLKAKLTI PTGGTMLKAKLTI PTGGTMLKAKLTI PTGGTMLKAKLTI PTGGTMLKAKLTI PTGGTMLKAKLTI PTGGTMLKAKLTI PTGGTMLKAKLTI PTGGTMLKAKLTI PTGGTMLKAKLTI PTGGTMLKAKLTI PTGGTMLKAKLTI PTGTMINNLQDKE P-LVHMINNLQDKE P-LVHMINNLQDKE P-LVHMINNLQDKE P-LVHMINNLQDTE |
| - | FYLLANT PRONHSK SYTA STATE TER SYTA STATE TO BE SYTA STATE TO BE SYTA STATE TO BE SYTA STATE TO BE SYTA STATE TO BE SYTA STATE TO STATE TO NA I I SY TO SERSH NA I I SY TONNEK SM KSI VQY TONNEK SM KSI VQY TONNEK SM KSI VQY TONNEK SM KSI VQY TONNEK SM | * 540 ** 520 ** 540 ** EALTHEILE BY THE TRILLE VLEALTHEILE BY THE TRILLE VLESHTE VLAR BENERKE IN HILLE BY THE TRILLE V VLESHTE VLAR BENERKE IN HILLE BY STANDAR LER VLESHTE VLAR BENERKE VRHILLE BY STANDAR LER I I SALTHE VLAR BY VRHILLE BY STANDAR LER I I SALTHE VLAR BY VRHILLE BY STANDAR LIT MLEARTHEILE BY STANDAR LER I I SALTHEILE STANDAR LER I I SALTHEILE STANDAR LER I I SALTHEILE STANDAR LER I I SALTHEILE STANDAR LER I I SALTHEILE STANDAR STEIMIR I VPGLEL SHASHTERIMIR I VPGLEL SHASHTERIMIR I VPGLEL SHASHTERIMIR I VPGLEL SHASHTERIMIR I VPGLEL SHASHTERIMIR I VPGLEL SHASHTERIMIR I VPGLEL SHASHTERIMIR I VPGLEL SHASHTERIMIR I VPGLEL SHASHTERIMIR I VPGLEL SHASHTERIMIR I VPFGLET SHASHTERIMIR I VPF |
| , | NAI NAI NAI NAI NAI NAI NAI NAI NAI NAI | |
| 01 | | + HARERER AFTER A TERM |
| ÷. | | MIFARY PETTANDARRO INFILLED VI BALLY PETTANDARRO INFILLED VI BENTHANDARRO INFILLED VI BENTHAND VI BENT |
| 1 | | 11ABNDKRRGY VLABNDKRRGY VLABNSKRHSE VLABNSKRHSE VLABNSKRHSE VLABNSKRUS IIABNSKRE SHASNTSKLE SHASNTSKRE SHASNTSKRE GHASNTSKRE GHASNTSKRE |
| ı | | S S N T T T S S N T T T T T T T T T T T |
| + ! i | ວາກວານ ທະຫາຍ ທະຫານ ຊື່ ຊື່ ຊື່ ຊື່ ຊື່ ຊື່ ຊື່ ຊື່ ຊື່ ຊື່ | SEO VIS VIS VIS VIS VIS VIS VIS VIS VIS VIS |
| | | |
| 1 | | |
| 2 | တီးထွားတိုးတိုးတိုးတိုးတိုးတိုးပြီး ဗို ဗို ဗို ဗို ဗို ဗို | |
| 7 | | RAPA RAPA RAPA RAPA RAPA RAPA RAPA RAPA |
| ļ | | |
| | ATTER KALLER KALLER KALLER KALLER RAVGI RAVGI RAVGI | |
| • | | |
| i | | |
| ì | | SOLOPPEN STANDS OF THE STANDS |
| 2 | | |
| ñ | 3 5 5 5 5 5 6 6 6 6 6 6 6 6 6 6 6 6 6 6 | |
| į | | |
| | O COCOCOCOCOCOCOCOCOCOCOCOCOCOCOCOCOCOC | |
| • | GVRIKEL GIRLLKT GIGUTKU GIGUTKI GIRLTRT GINVKRII GIKVKSII GIKVKSII GIKVKSII | 480 TERM TERM TERM TERM TERM TERM TERM TERM |
| | GORIN GORINI GORINI GORINI GORINI GORINI GORINI GORINI | 480 GHALTYOFDIBG GHELTYOFDIBG GHELTYOFDIBG GHELTYOFDIBG GHALTYOFDIBG GHALTYOFDIBG GHALTYOFDIBG CHALTYOFT CHALTY CHALTYOFT CHALTYOFT CHALTYOFT CHALTYOFT CHALTYOFT CHALTYOF |
| | DATE DATE TO THE CONTROL OF THE CONT | GGALTOL GGALTOL GGELTVOL GGELTVOL GGALTOL GGALTOL GGALTOL GGALTICOL GGALTICKI GGALTICKI GGALTICKI GGALTICKI |
| 09 | NET TO SEE TO SE | odatří odatří odatří odatří odatří odatří odatří |
| m | | |
| | | |
| _ | | |
| | | |
| | SPTRV3FG34GTGRDAKEPDITERTING TKSRL3FG34GTGREDVEFDITERTING SPTRL3FG34GTGREDAKEPDITERTING SPTRL3FG34GTGREDAKEPDITERTING SPTRL3FG34GTGREDVEFDITERTING SPTRV3FG34GTGREDVEFREDTGREN SSRR13L3ASSVARPGLRREDTGREN SSRR13L3ASSVARPGLREDDGREN SSRR13L3ASSVARPGLREDDGREN SSRR13L3ASSSVARPGLREDDGREN SSRR13L3ASSSVARPGLREDDGREN SSRR13L3ASSVARPGLREDDGREN SSRR13L3ASSSVARPGLREDDGREN SSRR13L3ASSSVARPGLREDDGREN SSRR13L3ASSSVARPGLREDDGREN SSRR13L3ASSSVARPGLREN SSRR13LSASSSVARPGLREN 3LSASSSVARP SSRR13LSASSSVARP SSRR13LSASSSVARP SSRR13LSASSSVARP SSRR13LSASSSVARP SSRR13LSASSSVARP SSRR13LSASSSVARP SSRR13LSASSSVARP SSRR13LSASSSVARP SSRR13LSASSSVARP SSRR13LSASSSVARP SSRR13LSASSVA | STPS-[STDANTIMPO SSAVNISHON SSAND SSAVNISHON SSAND KYT-NISHON SIMPI SSIPE STON WITHPO RYTINISHON STUPPI LEGDOLSAEV WEDLIND -SGUELSAEV WEDLIND -SGUELSAEV WEDLIND -LONELSAEV WEDLIND |
| | | |
| 340 | SPTRVE SEED OF TREATMENT OF THE SEED OF THE SEED OF THE SEED OF THE SEED OF THE SEED OF THE SEED OF THE SEED SEED SEED SEED SEED SEED SEED SE | BTPS-ESTEANT SSAVNESSEAS SSAVNESSEAS SSAVNESSEAS KYT-NESSEAS KYT-NESSEAS ENGLESASVAN SGDELSASVAN SGDELSASVAN - SGNELSASVAN - SGNELSASVAN - LUNELSANVAN - LUNELSANVAN - SGNELSANVAN - SGNELSANVAN - SGNELSANVAN - SGNELSANVAN |
| m | · [24 | BTPS-ESTER SSAVNESSER SSAVNESSER SSAVNESSER SSIPERSTER RYTINESSER - SGDELSAR - LONELSAR - LONELSAR - SGDELSAR - SGDELSAR - SGDELSAR - SGDELSAR - SGDELSAR - SGDELSAR |
| | SSONKI GEOLOGICAL | ANN AN AN AN AN AN AN AN AN AN AN AN AN |
| | | BTPS-ESTER SSAVNESHER SSAVNESHER SSAVNESHER SKIT-NESKER RYTINESKER LRGDQLSAEV - SGDELSAAV - LGNELSAAV - LGNELSAAV - LGNELSAAV |
| | | |
| | Potatos Arabidopsiss Adonissi Adonissi Tomatos Arabidopsiss Adoniss Peppers Tomatos Tomatos Adoniss | Potatos Arabidopsiss Adonissi Adonissi Tomatos Arabidopsiss Adoniss Tomatos Tomatos |
| | Potatos Arabidops Adonissi Adonissi Lettucess Tomatos Arabidops Adoniss Peppers Popacos Tomatos Tomatos Marigolds | Potatos Arabidops Adonissi Adonissi Lettucess Tomatos Arabidops Adoniss Tomatos Tomatos |
| | Potatos Arabidop Adonissi Lattuces Tomatos Arabidop Adoniss Peppers Poppers Tomatos Tomatos Marigold | PotatoB Arabidop Adonissi Iettucesi Tomatos Adoniss Tomatos Tomatos Tobaccos |
| | Are Are Too Too Too Too Too Too Too Too Too To | Pot Add Tob Tob Dar |
| | | |
| | | |

FIGURE 26

| | 103 102 102 102 100 84 | 57 208 207 207 212 205 194 | 167 318 317 317 322 315 | 276 422 427 427 427 431 414 | |
|-----|--|---|--|--|---|
| • | SKTAVDKEPPIS : AKLADKEPPIP : AKLADKEPPIP : SKLSEKKAQIP : SKLSEKRAQIS : SKLSEKRAQIS : | 220 RCVESCUVE RC | PUSPIRATE * PUSPIR | 440 NI: GOSSAVN: GOSSAVN: GOSSAVN: B-SSIPS: | |
| 100 | | • | 320 Barypishya Berypishya Berypishya Barypishya Barypishya Barypishya | 440 HREHTKQINSN HREDNSAYVSGQSSAVN HRODNSAYVSGQSSAVN HRODGSKENISLGKYT-N HRODGSKENISLGKYT-N HROHYSKNMLT8-SSIPSI | 540 CATLIRTITI CATHIKUTIKO CACHVRATIER CACHVRATIER CATLIRTITI |
| • | BILEVOM BILEVOM BILEVOM BILEVOM QLVEVOM | 200 TIDDDDP ILLIGRAYGRVSRHI TIDDDDR PILLIGRAYGRVSRHI TIDNDA PILLIGRAYGRVSRHI TIDNDA PILLIGRAYGRVSRHI TIDDDE PILLIGRAYGRVSRHE TIDDDE PILLIGRAYGRVSRHF | | A 420 A 5 VI A B A 6 VI A B A 6 V | ************************************** |
| 80 | SSO FI | ###### ###### | 300 PSLAVFADYRDYVRHDAQSL PDQAVFADYRDYTVRKVRSL PNLAVFADYRDYMQQKLQCS PNLAVFADYRDYMQCKLQCS PNLAVFADYRDYMQKLQCS PSLAVFADYRFFKKKPESL PSLAVFADYRFKKRDAGSL PSLAVFADYRFKKRSQSL | * 400 *LAFGAAASWYHPATGYSVVRSLSEAPKCA %LAFGAAASWYHPATGYSVVRSLSEAPKYA %LAFGAAASWYHPATGYSVVRSLSEAPKYA %LAFGAAASWYHPATGYSVVRSLSEAPKYA %LAFGAAASWYHPATGYSVVRSLSEAPKYA 10AFGAAASWYHPATGYSVVRSLSEAPKCA | 520 IIIABNDMBRGGI VISBNNIBRGGI VLABNSMBRSI VIABHSLENEDI VIABHSLENEDI IIABNDMRRGI |
| • | PPOVKRYSTRNIRFGL-CSV-RASGGGSGGRSCVAVREDFAD INLSSSKLAYNIHRYGSSCRVDPQVRADGGSGSRSSVAYKEGFAD INLSSSKLAYNIHRYGSSCRVDPQVRADGGSGSRTSVAYKEGFVD IHRFSLLKQRRFTNLSA-SSSLRQIKCSAKSDRCVVDKQGISVAD IMSGGELCQEKSIFLAY-EQYKSKCNSSSGSDSCVVDKEDFAD | 0 2 2 4 4 0 2 5 | EVEVDNAPFO EVEVENAPYD EVEVENAPYD EVEVENAPYD EVEVENAPYD EVEVENAPFO EVEVENAPFO | 400 ASKWIPATG ASKWIPATG ASKWIPATG ASKWIPATG ASKWIPATG ASKWIPATG | BULSALYNE EVLSSEYNE EVLSSEYNE EVLSSEYNE EVLSSEYNE EVLSSEYNE EVLSS |
| 60 | GGGSSGGSS ADGGSGSRT ADGGSGSRT CSAKSDR CNSSSGSDS | - DEFADICION NEDEFADICION NEDEFADICIER NEDEFADICIER NEDEFADICION NEDEFADICION | 280 CVQTAXGV CVQTAXGV CVQTAXGV CVQTAXGI SVQTAXGI | | 500 Gerigsslskadd Gerigsslskadd Wefigsslssed Wefigsslssed Wefigsslssed Wefigsslssed Wefigsslssed Wefigsslo |
| • | L-CSVRAS SSCRVDFQVH SSCRVDFQVH R-SSSLRQIK Y-BQYBSK | 160 | 9 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 | * ** ** ** ** ** ** ** ** ** | V PENNAN L PENNAN L PENNAN L PENNAN L PENNAN |
| 40 | PEPVYKRYSTRNIRFGL-CSVRASGGGSSGSESCVAVREDF INLSSSKLAYNIHRYGSSCRVDFQVRADGGSGSRSSVAYKEGF INLSSSKLAYNIHRYGSSCRVDFQVRADGGSGSRTSVAYKEGF HKFSLLKQRRFTNLSA-SSSLRQIKCSAKSDRCVVDKQGI MSGGGLCQEKSIFLAY-EQYRSKCNSSSGSDSCVVDKEDF | | 260 TVASGAASGKI TVASGAASGKI TVASGAASGKI TVASGAASGKI TVASGAASGKI TVASGAASGKI TVASGAASGKI | 1.GVRIKEIYEEEMSYIPVGGSLPNTEQKU 1.GIRLIKAYEEEMSYIPVGGSLPNTEQKU 1.GIQVAKAYEEEMSYIPVGGSLPNTEQKU 1.GIQVAKIYEEEMSYIPVGGSLPNTEQKU MGIRITRATYEEEMSYIPVGGSLPNTEQKU 1.GVRIKEIYEEEMSYIPVGGSLPNTEQKU | ALEST CONTROL OF CONTR |
| • | | 140 AGANESAKKGU SIRABAKKGU SIRABAKKGU AGANESAKKGU AGANESAKKGU AGANESAKKGU | | 360 RLMTLGVRIKEI RLDTLGIRILKT RLKTLGIQVTKY RLKTLGIQVTKI RLKANGIRITRT RLMTLGVRIKEI | ALITOPHOSE SELECTION OF SELECTI |
| 20 | AVSTEPSWS ISSOPWT ISSOPWT NVFTCPRFTD AVTTSPRIM | 120 DGALDHYVIGGGPAGIA ESVYDLYVIGGGPAGIS NCILDHYVIGGGPAGIS OTVIDHYVIGGGPAGIA | 240 IVENTROUS PROCEDUVI ITERATE SUVOENE IN ITERATE SUVOENDINI ITERATE SUVOENDINI ITERATE SUVOENDINI ITERATE SUVOENDINI ITERATE SUVOENDINI IVENTROSSI SUVOE | DULK KENE DULK K | 460 RORSFELFGL RORAFFLFGL RORAFFLFGL RORAFFLFGL RORAFFLFGL RORAFFLFGL |
| • | WECUGARNF - AAMAUSTFIJSWS - CRRK WELLGURNL ISSCIVIT - FOTR WECFGARNHTATHAVFTCIRFTDCNIR WECVGYQNV - GAMAULTRIRLN R SHRAG - HHTATMAAFTCIRFW | 120 IGBGALDHYVIGGGPAGLARANE FG | TANGERYDRIV BATACHSIN TESKYDRIV BATACHSIN TESKYDRIV BADDOCHSIN TESKYBRIT BADDOCHSIN TESKYBRIT BAPACHS BAPACHS BAPACHS BAPACHS BAPACHS BAPACHS BAPACHS BAPACHS BAPACHS BAPACHS BAPACHS BAPA | 340 FEETCLASKDAMPED FEETCLASKDAMPED FEETCLASKDAMPED FEETCLASKDAMPED FEETCLASKDAMPED FEETCLASKDAMPED FEETCLASKBAMPED | 460 STOAWNTUMPOERKRORSFELFGLALIUV SKOAWSTUMPERKRORAFFLFGLELIUV SKOAWSTUMPERKRORAFFLFGLELIUV SKOAWSTUMPERKRORAFFLFGLELIUV SROAWNTUMPOERKRORSFFLFGLALIUV STOAWNTUMPOERKRORSFFLFGLALIUV STOAWNTUMPOERKRORSFFLFGLALIUV |
| | Arabidopsiss : Adoniss1 : Adoniss2 : Lettucess : Tomatos : Marigolds : | Potatos Arabidopsiss: Adoniss: Adoniss: Lettucess: Tomatos: Marigolds: | Potatos Arabidopsiss : Adonissi : Adonissi : Lettucess : Tomatos : | Potatos : Arabidopsiss : Adonissi : Lettucess : Tomstos : Marigolds : | Potatos : Arabidopsiss : Adonissi : Adonissi : Lettucess : Tomatos : Marigolds : |
| | | | | | |



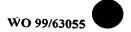


FIGURE 28

GAP of: Arabidopsis epsilon cyclase to Lettuce epsilon cyclase

Gap Weight: 12 Average Match: 2.912
Length Weight: 4 Average Mismatch: -2.003

Quality: 1837 Ratio: 3.499 Length: 534

Percent Similarity: 76.381 Percent Identity: 69.905 Gaps:

Match display thresholds for the alignment(s):

= IDENTITY

: = 2

1

Arabidopsis x Lettuce

| 1 MECVGARNF.AAMAVSTFPSWSCRRKFPVVKRYSYRNIRFGLCSVRA 46 |
|---|
| 1 MECFGARNMTATMAVFTCPRFTDCNIRHKFSLLKQRRFTNLSASSSLRQI 50 |
| 47 SGGGSSGSESCVAVREDFADEEDFVKAGGSEILFVQMQQNKDMDEQSKLV 96 |
| 51 KCSAKSDRCVVDKQGISVADEEDYVKAGGSELFFVQMQRTKSMESQSKLS 100 |
| 97 DKLPPISIGDGALDHVVIGCGPAGLALAAESAKLGLKVGLIGPDLPFTNN 146 |
| : . |
| 147 YGVWEDEFNDLGLQKCIEHVWRETIVYLDDDKPITIGRAYGRVSRRLLHE 196 |
| : : :: : |
| 197 ELLRRCVESGVSYLSSKVDSITEASDGLPLVAGDDVRVIV |
| |
| 247 ASGKLLQYEVGGPRVCVQTAYGVEVEVENSPYDPDQMVFMDYRDYTNEKV 296 |
| : . |
| |
| 297 RSLEAEYPTFLYAMPMTKSRLFFEETCLASKDVMPFDLLKTKLMLRLDTL 346 . :: . :: :: :: ::: : |
| 350 |
| 347 GIRILKTYEEEWSYIPVGGSLPNTEQKNLAFGAAASMVHPATGYSVVRSL 396 : |
| 400 |
| 397 SEAPKYASVIAEILREETTKQINSNISRQAWDTLWPPERKRQRAF 441 |
| |
| |
| 442 FLFGLALIVQFDTEGIRSFFRTFFRLPKWMWQGFLGSTLTSGDLVLFALY 491 |
| |
| 492 MFVISPNNLRKGLINHLISDPTGATMIKTYLKV* 525 |
| . . : : : : 501 MFVIAPHSLRMELVRHLLSDPTGATMVKAYLTI* 534 |



SEQUENCE LISTING

| <110> | CUNNII | | | , FF | ANCI | s x. | | | | | | | | | |
|---------------------------|----------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|-----|
| <120> | GENES METHOI | | | | | | THES | sis <i>p</i> | 4 dna | /ETAF | BOLIS | IA M | ND | | |
| <130> | 8172-9 | 9023 | | | | | | | | | | | | • | |
| <140> <141> | | | | ED . | | | | ٠ | | | | | | • | |
| <150> <151> | | | | | | | | | , | | | | | | |
| <150> <151> | | | | | | | | | | | | | | | |
| <160> | 61 | | | | | | | | | | | | | | |
| <170> | Patent | :In V | er. | 2.0 | | | | | | | | | | • | |
| <210><211><211><212><213> | 1860 DNA | dopsi | ls th | nalia | ına | | | | | | | | | | |
| <220> <221> <222> | | (16 | 580) | | | | | | | | | | • | | |
| <400> ACAAAA | | raat <i>e</i> | TAG | AT TO | CTCI | rttci | r GCT | rtgct | ATA | CCTI | · 『GATA | AGA A | ACAAT | TATAAC | 60 |
| Aatggi | 'GTAA (| STCTI | CTC | SC TO | TAT | rcga <i>i</i> | A ATI | ratt1 | rgga | GGAG | GAA. | Met | | TGT Cys | 117 |
| GTT GG Val Gl | G GCT y Ala 5 | AGG Arg | AAT Asn | TTC Phe | GCA Ala 10 | GCA Ala | ATG Met | GCG Ala | GTT Val | TCA Ser 15 | ACA Thr | TTT Phe | CCG Pro | TCA Ser | 165 |
| rgg Ag Frp Se 20 | T TGT r Cys | CGA Arg | AGG Arg | AAA Lys 25 | TTT Phe | CCA Pro | GTG Val | GTT Val | AAG Lys 30 | AGA Arg | TAC Tyr | AGC Ser | TAT Tyr | AGG Arg 35 | 213 |
| AAT AI Asn Il | T CGT e Arg | TTC | GGT Gly 40 | TTG Leu | TGT Cys | AGT Ser | GTC Val | AGA Arg 45 | GCT Ala | AGC Ser | GGC Gly | GGC Gly | GGA Gly 50 | AGT Ser | 261 |
| TCC GG Ser Gl | T AGT y Ser | GAG Glu 55 | AGT Ser | TGT Cys | GTA Val | GCG Ala | GTG Val 60 | AGA Arg | GAA Glu | GAT Asp | TTC Phe | GCT Ala 65 | GAC Asp | GAA Glu | 309 |
| GAA GA Glu As | T TTT p Phe 70 | GTG Val | AAA Lys | GCT Ala | GGT Gly | GGT Gly 75 | TCT Ser | GAG Glu | ATT Ile | CTA Leu | TTT Phe 80 | GTT Val | CAA Gln | ATG Met | 357 |
| | | | | | | | | | | | 00 | | | | |

| | W | 99/6 | 3055 | | | | | | | | | | | | | PCT/US99/12121 |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|----------------|
| Gln | Gln 85 | Asn | Lys | Asp | Met | Asp 90 | Glu | Gln | Ser | Lys | Leu 95 | Val | Asp | Lys | Leu | - * - |
| CCT Pro 100 | Pro | ATA Ile | TCA Ser | ATT Ile | GGT Gly 105 | GAT Asp | GGT Gly | GCT Ala | TTG Leu | GAT Asp 110 | CAT His | GTG Val | GTT Val | ATT Ile | GGT Gly 115 | 453 . |
| TGT Cys | GGT Gly | CCT Pro | GCT Ala | GGT Gly 120 | TTA Leu | GCC Ala | TTG Leu | GCT Ala | GCA Ala 125 | GAA Glu | TCA Ser | GCT Ala | AAG Lys | CTT Leu 130 | GGA Gly | 501 |
| TTA Leu | AAA Lys | GTT Val | GGA Gly 135 | CTC Leu | ATT Ile | GGT Gly | CCA Pro | GAT Asp 140 | CTT Leu | CCT Pro | TTT Phe | ACT Thr | AAC Asn 145 | AAT Asn | TAC Tyr | 549 |
| GGT Gly | GTT Val | TGG Trp 150 | GAA Glu | GAT Asp | GAA Glu | TTC Phe | AAT Asn 155 | GAT Asp | CTT Leu | GGG Gly | CTG Leu | CAA Gln 160 | AAA Lys | TGT Cys | ATT Ile | 597 |
| GAG Glu | CAT His 165 | GTT Val | TGG Trp | AGA Arg | GAG Glu | ACT Thr 170 | ATT Ile | GTG Val | TAT Tyr | CTG Leu | GAT Asp 175 | GAT Asp | GAC Asp | AAG Lys | CCT Pro | 645 |
| ATT Ile 180 | ACC Thr | ATT Ile | GGC Gly | CGT Arg | GCT Ala 185 | TAT Tyr | GGA Gly | AGA Arg | GTT Val | AGT Ser 190 | CGA Arg | CGT Arg | TTG Leu | CTC Leu | CAT His 195 | 693 |
| GAG Glu | GAG Glu | CTT Leu | TTG Leu | AGG Arg 200 | AGG Arg | TGT Cys | GTC Val | GAG Glu | TCA Ser 205 | GGT Gly | GTC Val | TCG Ser | TAC Tyr | CTT Leu 210 | AGC Ser | 741 |
| TCG Ser | AAA Lys | GTT Val | GAC Asp 215 | AGC Ser | ATA Ile | ACA Thr | GAA Glu | GCT Ala 220 | TCT Ser | GAT Asp | GGC Gly | CTT Leu | AGA Arg 225 | CTT Leu | GTT Val | 789 |
| GCT Ala | TGT Cys | GAC Asp 230 | GAC Asp | AAT Asn | AAC Asn | GTC Val | ATT Ile 235 | CCC Pro | TGC Cys | AGG Arg | CTT Leu | GCC Ala 240 | ACT Thr | GTT Val | GCT Ala | 837 |
| TCT Ser | GGA Gly 245 | GCA Ala | GCT Ala | TCG Ser | GGA Gly | AAG Lys 250 | CTC Leu | TTG Leu | CAA Gln | TAC Tyr | GAA Glu 255 | GTT Val | GGT Gly | GGA Gly | CCT Pro | 885 |
| AGA Arg 260 | GTC Val | TGT Cys | GTG Val | CAA Gln | ACT Thr 265 | GCA Ala | TAC Tyr | GGC Gly | GTG Val | GAG Glu 270 | GTT Val | GAG Glu | GTG Val | GAA Glu | AAT Asn 275 | 933 |
| AGT Ser | CCA Pro | TAT Tyr | GAT Asp | CCA Pro 280 | GAT Asp | CAA Gln | ATG Met | GTT Val | TTC Phe 285 | ATG Met | GAT Asp | TAC Tyr | AGA Arg | GAT Asp 290 | TAT Tyr | 981 |
| ACT Thr | AAC Asn | GAG Glu | AAA Lys 295 | GTT Val | CGG Arg | AGC Ser | TTA Leu | GAA Glu 300 | GCT Ala | GAG Glu | TAT Tyr | CCA Pro | ACG Thr 305 | TTT Phe | CTG Leu | 1029 |
| TAC Tyr | GCC Ala | ATG Met 310 | CCT Pro | ATG Met | ACA Thr | AAG Lys | TCA Ser 315 | AGA Arg | CTC Leu | TTC Phe | TTC Phe | GAG Glu 320 | GAG Glu | ACA Thr | TGT Cys | 1077 |

1125

TTG GCC TCA AAA GAT GTC ATG CCC TTT GAT TTG CTA AAA ACG AAG CTC Leu Ala Ser Lys Asp Val Met Pro Phe Asp Leu Leu Lys Thr Lys Leu

PCT/US99/12121

÷ ...

| - | | | |
|---|--|--|--|
| | | | |

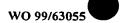
WO 99/63055

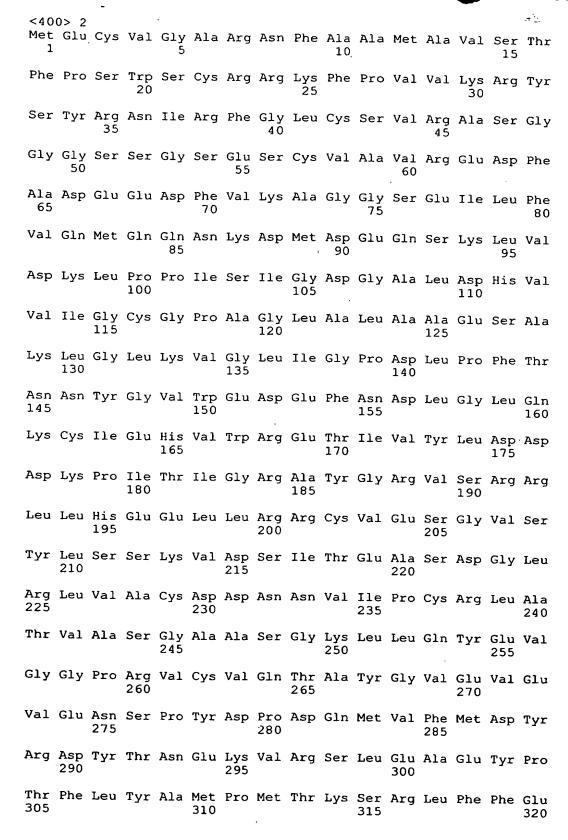
| | 325 | | | | | 330 | | | | | 335 | | | | | |
|-----|------|-----|------|-------|-------------------|--------|------|------|------|-------|-----|-------|------|------|-------------------|------|
| | | | | | ACA Thr 345 | | | | | | | | | | | 1173 |
| | | | | | ATC Ile | | | | | | | | | | | 1221 |
| | | | | | TTT Phe | | | | | | | | | | | 1269 |
| | | | | | GTG Val | | | | | | | | | | | 1317 |
| | | | | | ATA Ile | | | | | | | | | | | 1365 |
| | | | | | CAA Gln 425 | | | | | | | | | | - | 1413 |
| | | | | | TTC Phe | | | | | | | | | | | 1461 |
| | | | | Gly | ATT Ile | | | | | | | | | | | 1509 |
| | | | Met | | CAA Gln | | | | | | | | Thr | | GGA Gly | 1557 |
| | | Val | | | GCT Ala | | Tyr | | | | | Ser | | | AAT Asn | 1605 |
| | Arg | | | | | Asn | | | | | Asp | | | | GCA Ala 515 | 1653 |
| | | | | | TAT Tyr | | | | | TTTA | CTT | ATCA | ACTC | TT | | 1700 |
| AGG | TTTG | TGT | ATAT | TATAT | GT T | 'GAT'I | TATC | T GA | ATAA | TCGA | TCA | AAGA | ATG | GTAT | GTGGGT | 1760 |
| TAC | TAGG | AAG | TTGG | SAAAC | AA A | CATO | TATA | G AA | TCTA | AGGA | GTO | SATCO | AAA | TGGA | GATGGA | 1820 |
| AAC | GAAA | AGA | AAAA | AATC | CAG I | CTTI | GTTI | T GI | GGTI | 'AGTG | i | | | | | 1860 |

<210> 2

<211> 524 <212> PRT

<213> Arabidopsis thaliana





PCT/US99/12121

480



| | WU | 99/03 | 0033 | , | | | | | | | | | | | | rC1/0 | 377/1 | 411 |
|------------|---------------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-------|-------|-----|
| Glu | Thr | Cys | Leu | Ala 325 | Ser | Lys | Asp | Val | Met 330 | Pro | Phe | Asp | Leu | Leu 335 | Lys | | | - |
| Thr | Lys | Leu | Met 340 | Leu | Arg | Leu | Asp | Thr 345 | Leu | Gly | Ile | Arg | Ile 350 | Leu | Lys | | | |
| Thr | Tyr | Glu 355 | Glu | Glu | Trp | Ser | Tyr 360 | Ile | Pro | Val | Gly | Gly 365 | Ser | Leu | Pro | | | |
| Asn | Thr 370 | Glu | Gln | Lys | Asn | Leu 375 | Ala | Phe | Gly | Ala | Ala 380 | Ala | Ser | Met | Val | | | |
| His 385 | Pro | Ala | Thr | Gly | Tyr 390 | Ser | Val | Val | Arg | Ser 395 | Leu | Ser | Glu | Ala | Pro 400 | | | |
| Lys | Tyr | Ala | Ser | Val 405 | Ile | Ala | Glu | Ile | Leu 410 | Arg | Glu | Glu | Thr | Thr 415 | Lys | | | |
| Gln | Ile | Asn | Ser 420 | Asn | Ile | Ser | Arg | Gln 425 | Ala | Trp | Asp | Thr | Leu 430 | Trp | Pro | | | |
| Pro | Glu | Arg 435 | Lys | Arg | Gln | Arg | Ala 440 | Phe | Phe | Leu | Phe | Gly 445 | Leu | Ala | Leu | | | |
| Ile | Val 450 | Gln | Phe | Asp | Thr | Glu 455 | Gly | Ile | Arg | Ser | Phe 460 | Phe | Arg | Thr | Phe | | | |
| Phe 465 | Arg | Leu | Pro | Lys | Trp 470 | Met | Trp | Gln | Gly | Phe 475 | Leu | Gly | Ser | Thr | Leu 480 | | | |
| Thr | Ser | Gly | Asp | Leu 485 | Val | Leu | Phe | Ala | Leu 490 | Tyr | Met | Phe | Val | Ile 495 | | | | |
| Pro | Asn | Asn | Leu 500 | Arg | Lys | Gly | Leu | 11e 505 | Asn | His | Leu | Ile | Ser 510 | Asp | Pro | | | |
| Thr | Gly | Ala 515 | Thr | Met | Ile | Lys | Thr 520 | Tyr | Leu | Lys | Val | | | | | | | |
| <213 | 0> 3 1> 9: 2> Di 3> A: | | dops | is t | hali | ana | | | | | | | | | | | | |
| _ | 0> 3 C TTT (| CTC (| CTCC | TCCT | CT A | CCGA | TTTC | C GA | CTCC | GCCT | CCC | GAAA | TCC | TTAT | CCGG | ΑT | 6 | 0 |
| TCT | CTCC | GTC | TCTT | CGAT | TT A | AACG | CTTT | т ст | GTCT | GTTA | CGT | CGTC | GAA | GAAC | GGAG | AC | 12 | 0 |
| AGA | ATTC' | TCC | GATT | GAGA | AC G | ATGA | GAGA | c cg | GAGA | GCAC | GAG | СТСС | ACA | AACG | CTAT | AG | 18 | 0 |
| ACG | CTGA | GTA | TCTG | GCGT | TG C | GTTT | GGCG | G AG | TAAA | TGGA | GAG | GAAG | AAA | TCGG | AGAG | GT . | 24 | 0 |
| CCA | CTTA | TCT | AATC | GCTG | CT A | TGTT | GTCG | A GC | TTTG | GTAT | CAC | TTCT | ATG | GCTG | TTAT | GG | 30 | 0 |
| CTG | TTTA | CTA | CAGA | TTCT | CT T | GGCA | AATG | G AG | GGAG | GTGA | GAT | CTCA | ATG | TTGG. | TAAA | GT | 36 | 0 |
| TTG | GTAC | ATT | TGCT | CTCT | CT G | TTGG | TGCT | G CT | GTTG | GTAT | GGA | ATTC | TGG | GCAA | GATG | GG | 42 | 0 |

CTCATAGAGC TCTGTGGCAC GCTTCTCTAT GGAATATGCA TGAGTCACAT CACAAACCAA

WO 99/63055



PCT/US99/12121

| GAGAAGGACC GTTTGAGC | TA AACGATGTT | T TTGCTATAGT | GAACGCTGGT CCAGCGATTG | | | | | | | |
|---|--------------------|--------------------|----------------------------|--|--|--|--|--|--|--|
| GTCTCCTCTC TTATGGAT | тс ттсаатааа | G GACTCGTTCC | TGGTCTCTGC TTTGGCGCCG | | | | | | | |
| GGTTAGGCAT AACGGTGT | TT GGAATCGCC | T ACATGTTTGT | CCACGATGGT CTCGTGCACA | | | | | | | |
| AGCGTTTCCC TGTAGGTC | CC ATCGCCGAC | G TCCCTTACCT | CCGAAAGGTC GCCGCCGCTC | | | | | | | |
| ACCAGCTACA TCACACAG | AC AAGTTCAAT | G GTGTACCATA | TGGACTGTTT CTTGGACCCA | | | | | | | |
| AGGAATTGGA AGAAGTTG | GA GGAAATGAA | G AGTTAGATAA | GGAGATTAGT CGGAGAATCA | | | | | | | |
| AATCATACAA AAAGGCCT | CG GGCTCCGGG | T CGAGTTCGAG | TTCTTGACTT TAAACAAGTT | | | | | | | |
| TTAAATCCCA AATTCTTT | TT TTGTCTTCT | G TCATTATGAT | CATCTTAAGA CGGTCT | | | | | | | |
| TTAAATCCCA AATTCTTTT TTGTCTTCTG TCATTATGAT CATCTTAAGA CGGTCT <210> 4 <211> 294 <212> PRT <213> Arabidopsis thaliana | | | | | | | | | | |
| <400> 4 Ser Phe Ser Ser Ser | Ser Thr Asn | Phe Arg Iou | Arg Leu Pro Lys Ser | | | | | | | |
| 1 5 | ber im Asp | 10 | 15 | | | | | | | |
| Leu Ser Gly Phe Ser 20 | Pro Ser Leu | Arg Phe Lys 25 | Arg Phe Ser Val Cys 30 | | | | | | | |
| Tyr Val Val Glu Glu 35 | Arg Arg Gln 40 | Asn Ser Pro | Ile Glu Asn Asp Glu 45 | | | | | | | |
| Arg Pro Glu Ser Thr 50 | Ser Ser Thr 55 | Asn Ala Ile | Asp Ala Glu Tyr Leu 60 | | | | | | | |
| Ala Leu Arg Leu Ala 65 | Glu Lys Leu 70 | Glu Arg Lys 75 | Lys Ser Glu Arg Ser 80 | | | | | | | |
| Thr Tyr Leu Ile Ala 85 | Ala Met Leu | Ser Ser Phe 90 | Gly Ile Thr Ser Met 95 | | | | | | | |
| Ala Val Met Ala Val 100 | Tyr Tyr Arg | Phe Ser Trp 105 | Gln Met Glu Gly Gly 110 | | | | | | | |
| Glu Ile Ser Met Leu 115 | Glu Met Phe 120 | Gly Thr Phe | Ala Leu Ser Val Gly 125 | | | | | | | |
| Ala Ala Val Gly Met 130 | Glu Phe Trp 135 | Ala Arg Trp | Ala His Arg Ala Leu 140 | | | | | | | |
| Trp His Ala Ser Leu 145 | Trp Met Asn 150 | His Glu Ser 155 | His His Lys Pro Arg 160 | | | | | | | |
| Glu Gly Pro Phe Glu 165 | Leu Asn Asp | Val Phe Ala 170 | Ile Val Asn Ala Gly 175 | | | | | | | |
| Pro Ala Ile Gly Leu 180 | Leu Ser Tyr | Gly Phe Phe 185 | Asn Lys Gly Leu Val 190 | | | | | | | |
| Pro Gly Leu Cys Phe 195 | Gly Ala Gly 200 | Leu Gly Ile | Thr Val Phe Gly Ile 205 | | | | | | | |

WO 99/63055

Ala Tyr Met Phe Val His Asp Gly Leu Val His Lys Arg Phe Pro Val 210 215 220

Gly Pro Ile Ala Asp Val Pro Tyr Leu Arg Lys Val Ala Ala Ala His 225 230 235 240

Gln Leu His His Thr Asp Lys Phe Asn Gly Val Pro Tyr Gly Leu Phe 245 250 255

Leu Gly Pro Lys Glu Leu Glu Glu Val Gly Gly Asn Glu Glu Leu Asp 260 265 270

Lys Glu Ile Ser Arg Ile Lys Ser Tyr Lys Lys Ala Ser Gly Ser 275 280 285

Gly Ser Ser Ser Ser Ser 290

<210> 5

<211> 162

<212> PRT

<213> Alicalgenes sp.

<400> 5

Met Thr Gln Phe Leu Ile Val Val Ala Thr Val Leu Val Met Glu Leu 1 5 10 15

Thr Ala Tyr Ser Val His Arg Trp Ile Met His Gly Pro Leu Gly Trp 20 25 30

Gly Trp His Lys Ser His His Glu Glu His Asp His Ala Leu Glu Lys 35 40 45

Asn Asp Leu Tyr Gly Val Val Phe Ala Val Leu Ala Thr Ile Leu Phe 50 55 60

Thr Val Gly Ala Tyr Trp Trp Pro Val Leu Trp Trp Ile Ala Leu Gly 65 70 75 80

Met Thr Val Tyr Gly Leu Ile Tyr Phe Ile Leu His Asp Gly Leu Val 85 90 95

His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Arg Gly Tyr Phe Arg 100 105 110

Arg Leu Tyr Gln Ala His Arg Leu His His Ala Val Glu Gly Arg Asp 115 120 125

His Cys Val Ser Phe Gly Phe Ile Tyr Ala Pro Pro Val Asp Lys Leu 130 135 140

Lys Gln Asp Leu Lys Arg Ser Gly Val Leu Arg Pro Gln Asp Glu Arg 145 150 155 160

Pro Ser

<210> 6

<211> 175

<212> PRT

<213> Erwinia herbicola

<400> 6 Met Leu Asn Ser Leu Ile Val Ile Leu Ser Val Ile Ala Met Glu Gly Ile Ala Ala Phe Thr His Arg Tyr Ile Met His Gly Trp Gly Trp Arg Trp His Glu Ser His His Thr Pro Arg Lys Gly Val Phe Glu Leu Asn Asp Leu Phe Ala Val Val Phe Ala Gly Val Ala Ile Ala Leu Ile Ala Val Gly Thr Ala Gly Val Trp Pro Leu Gln Trp Ile Gly Cys Gly Met Thr Val Tyr Gly Leu Leu Tyr Phe Leu Val His Asp Gly Leu Val His Gln Arg Trp Pro Phe His Trp Ile Pro Arg Arg Gly Tyr Leu Lys Arg 105 Leu Tyr Val Ala His Arg Leu His His Ala Val Arg Gly Arg Glu Gly 120 Cys Val Ser Phe Gly Phe Ile Tyr Ala Arg Lys Pro Ala Asp Leu Gln 135 Ala Ile Leu Arg Glu Arg His Gly Arg Pro Pro Lys Arg Asp Ala Ala 155 Lys Asp Arg Pro Asp Ala Ala Ser Pro Ser Ser Ser Pro Glu <210> 7 <211> 175 <212> PRT <213> Erwinia uredovora <400> 7 Met Leu Trp Ile Trp Asn Ala Leu Ile Val Phe Val Thr Val Ile Gly Met Glu Val Ile Ala Ala Leu Ala His Lys Tyr Ile Met His Gly Trp 25 Gly Trp Gly Trp His Leu Ser His His Glu Pro Arg Lys Gly Ala Phe Glu Val Asn Asp Leu Tyr Ala Val Val Phe Ala Ala Leu Ser Ile Leu

Leu Ile Tyr Leu Gly Ser Thr Gly Met Trp Pro Leu Gln Trp Ile Gly

Ala Gly Met Thr Ala Tyr Gly Leu Leu Tyr Phe Met Val His Asp Gly

Leu Val His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr

100

PCT/US99/12121

WO 99/63055



Leu Lys Arg Leu Tyr Met Ala His Arg Met His His Ala Val Arg Gly
115 120 125

Lys Glu Gly Cys Val Ser Phe Gly Phe Leu Tyr Ala Pro Pro Leu Ser 130 135 140

Lys Leu Gln Ala Thr Leu Arg Glu Arg His Gly Ala Arg Ala Gly Ala 145 150 155 160

Ala Arg Asp Ala Gln Gly Gly Glu Asp Glu Pro Ala Ser Gly Lys 165 170 175

<210> 8

<211> 162

<212> PRT

<213> Agrobacterium aurianticum

<400> 8

Met Thr Asn Phe Leu Ile Val Val Ala Thr Val Leu Val Met Glu Leu 1 5 10 15

Thr Ala Tyr Ser Val His Arg Trp Ile Met His Gly Pro Leu Gly Trp
20 25 30

Gly Trp His Lys Ser His His Glu Glu His Asp His Ala Leu Glu Lys 35 40 45

Asn Asp Leu Tyr Gly Leu Val Phe Ala Val Ile Ala Thr Val Leu Phe 50 55 60

Thr Val Gly Trp Ile Trp Ala Pro Val Leu Trp Trp Ile Ala Leu Gly 65 70 75 80

Met Thr Val Tyr Gly Leu Ile Tyr Phe Val Leu His Asp Gly Leu Val 85 90 95

His Trp Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr Ala Arg
100 105 110

Arg Leu Tyr Gln Ala His Arg Leu His His Ala Val Glu Gly Arg Asp 115 120 125

His Cys Val Ser Phe Gly Phe Ile Tyr Ala Pro Pro Val Asp Lys Leu 130 135 140

Lys Gln Asp Leu Lys Met Ser Gly Val Leu Arg Ala Glu Ala Gln Glu 145 150 155 160

Arg Thr

<210> 9

<211> 954

<212> DNA

<213> Arabidopsis thaliana

<400> 9

CCACGGGTCC GCCTCCCCGT TTTTTTCCGA TCCGATCTCC GGTGCCGAGG ACTCAGCTGT

TTGTTCGCGC TTTCTCAGCC GTCACCATGA CCGATTCTAA CGATGCTGGA ATGGATGCTG

| WO 99/63055 | PCT/US99/12121 |
|-------------|----------------|
| | |

| • | | | | | | |
|--|-------------|------------|------------|------------|------------|-----|
| TTCAGAGACG A | ACTCATGTTT | GAAGACGAAT | GCATTCTCGT | TGATGAAAAT | AATCGTGTGG | 180 |
| TGGGACATGA C | CACTAAGTAT | AACTGTCATC | TGATGGAAAA | GATTGAAGCT | GAGAATTTAC | 240 |
| TTCACAGAGC T | TTTCAGTGTG | TTTTTATTCA | ACTCCAAGTA | TGAGTTGCTT | CTCCAGCAAC | 300 |
| GGTCAAAAAC A | AAAGGTTACT | TTCCCACTTG | TGTGGACAAA | CACTTGTTGC | AGCCATCCTC | 360 |
| TTTACCGTGA A | ATCCGAGCTT | ATTGAAGAGA | ATGTGCTTGG | TGTAAGAAAT | GCCGCACAAA | 420 |
| GGAAGCTTTT (| CGATGAGCTC | GGTATTGTAG | CAGAAGATGT | ACCAGTCGAT | GAGTTCACTC | 480 |
| CCTTGGGACG C | CATGCTTTAC | AAGGCACCTT | CTGATGGGAA | ATGGGGAGAG | CACGAAGTTG | 540 |
| ACTATCTACT C | CTTCATCGTG | CGGGATGTGA | AGCTTCAACC | AAACCCAGAT | GAAGTGGCTG | 600 |
| AGATCAAGTA C | CGTGAGCAGG | GAAGAGCTTA | AGGAGCTGGT | GAAGAAAGCA | GATGCTGGCG | 660 |
| ATGAAGCTGT C | GAAACTATCT | CCATGGTTCA | GATTGGTGGT | GGATAATTTC | TTGATGAAGT | 720 |
| GGTGGGATCA T | TGTTGAGAAA | GGAACTATCA | CTGAAGCTGC | AGACATGAAA | ACCATTCACA | 780 |
| AGCTCTGAAC T | TTTCCATAAG | TTTTGGATCT | TCCCCTTCCC | ТААТААТА | TAAGAGATGA | 840 |
| GACTTTTATT O | GATTACAGAC | AAAACTGGCA | ACAAAATCTA | TTCCTAGGAT | TTTTTTTGC | 900 |
| TTTTTATTTA C | CTTTTGATTC | ATCTCTAGTT | TAGTTTTCAT | СТТАААААА | AAAA | 954 |
| <210> 10 <211> 996 <212> DNA <213> Arabic | dopsis thal | iana | | | | |
| <400> 10 CACCAATGTC T | TGTTTCTTCT | ТТАТТТААТС | TCCCATTGAT | TCGCCTCAGA | TCTCTCGCTC | 60 |
| TTTCGTCTTC T | TTTTTCTTCT | TTCCGATTTG | CCCATCGTCC | TCTGTCATCG | ATTTCACCGA | 120 |
| GAAAGTTACC G | SAATTTTCGT | GCTTTCTCTG | GTACCGCTAT | GACAGATACT | AAAGATGCTG | 180 |
| GTATGGATGC T | TGTTCAGAGA | CGTCTCATGT | TTGAGGATGA | ATGCATTCTT | GTTGATGAAA | 240 |
| CTGATCGTGT T | TGTGGGGCAT | GTCAGCAAGT | ATAATTGTCA | TCTGATGGAA | AATATTGAAG | 300 |
| CCAAGAATTT G | GCTGCACAGG | GCTTTTAGTG | TATTTTTATT | CAACTCGAAG | TATGAGTTGC | 360 |
| TTCTCCAGCA F | AAGGTCAAAC | ACAAAGGTTA | CGTTCCCTCT | AGTGTGGACT | AACACTTGTT | 420 |
| GCAGCCATCC T | CTTTACCGT | GAATCAGAGC | TTATCCAGGA | CAATGCACTA | GGTGTGAGGA | 480 |
| ATGCTGCACA A | AAGAAAGCTT | CTCGATGAGC | TTGGTATTGT | AGCTGAAGAT | GTACCAGTCG | 540 |
| | | | | | | |

ATGAGTTCAC TCCCTTGGGA CGTATGCTGT ACAAGGCTCC TTCTGATGGC AAATGGGGAG

AGCATGAACT TGATTACTTG CTCTTCATCG TGCGAGACGT GAAGGTTCAA CCAAACCCAG

ATGAAGTAGC TGAGATCAAG TATGTGAGCC GGGAAGAGCT GAAGGAGCTG GTGAAGAAAG

CAGATGCAGG TGAGGAAGGT TTGAAACTGT CACCATGGTT CAGATTGGTG GTGGACAATT

600

660

720

WO 99/63055



PCT/US99/12121

| TCTTGATGAA | GTGGTGGGAT | CATGTTGAGA | AAGGAACTTT | GGTTGAAGCT | ATAGACATGA | 840 |
|------------|------------|------------|------------|------------|------------|-----|
| AAACCATCCA | CAAACTCTGA | ACATCTTTTT | TTAAAGTTTT | TAAATCAATC | AACTTTCTCT | 900 |
| TCATCATTTT | TATCTTTTCG | ATGATAATAA | TTTGGGATAT | GTGAGACACT | TACAAAACTT | 960 |
| CCAAGCACCT | CAGGCAATAA | TAAAGTTTGC | GGCCGC | • | | 996 |

| <210> 11 | |
|----------|--|
|----------|--|

<213> Haematococcus pluvialis

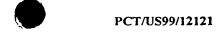
| < | Δ | n | n | > | 1 | 1 |
|---|---|---|---|---|---|---|
| • | 7 | v | v | _ | | _ |

| <400> 11 CTCGGTAGCT GGCCACAATC | GCTATTTGGA | ACCTGGCCCG | GCGGCAGTCC | GATGCCGCGA | 60 |
|-----------------------------------|------------|------------|------------|------------|------|
| TGCTTCGTTC GTTGCTCAGA | GGCCTCACGC | ATATCCCCCG | CGTGAACTCC | GCCCAGCAGC | 120 |
| CCAGCTGTGC ACACGCGCGA | CTCCAGTTTA | AGCTCAGGAG | CATGCAGATG | ACGCTCATGC | 180 |
| AGCCCAGCAT CTCAGCCAAT | CTGTCGCGCG | CCGAGGACCG | CACAGACCAC | ATGAGGGGTG | 240 |
| CAAGCACCTG GGCAGGCGGG | CAGTCGCAGG | ATGAGCTGAT | GCTGAAGGAC | GAGTGCATCT | 300 |
| TGGTGGATGT TGAGGACAAC | ATCACAGGCC | ATGCCAGCAA | GCTGGAGTGT | CACAAGTTCC | 360 |
| TACCACATCA GCCTGCAGGC | CTGCTGCACC | GGGCCTTCTC | TGTGTTCCTG | TTTGACGATC | 420 |
| AGGGGCGACT GCTGCTGCAA | CAGCGTGCAC | GCTCAAAAAT | CACCTTCCCA | AGTGTGTGGA | 480 |
| CGAACACCTG CTGCAGCCAC | CCTTTACATG | GGCAGACCCC | AGATGAGGTG | GACCAACTAA | 540 |
| GCCAGGTGGC CGACGGAACA | GTACCTGGCG | CAAAGGCTGC | TGCCATCCGC | AAGTTGGAGC | 600 |
| ACGAGCTGGG GATACCAGCG | CACCAGCTGC | CGGCAAGCGC | GTTTCGCTTC | CTCACGCGTT | 660 |
| TGCACTACTG TGCCGCGGAC | GTGCAGCCAG | CTGCGACACA | ATCAGCGCTC | TGGGGCGAGC | 720 |
| ACGAAATGGA CTACATCTTG | TTCATCCGGG | CCAACGTCAC | CTTGGCGCCC | AACCCTGACG | 780 |
| AGGTGGACGA AGTCAGGTAC | GTGACGCAAG | AGGAGCTGCG | GCAGATGATG | CAGCCGGACA | 840 |
| ACGGGCTGCA ATGGTCGCCG | TGGTTTCGCA | TCATCGCCGC | GCGCTTCCTT | GAGCGTTGGT | 900 |
| GGGCTGACCT GGACGCGGCC | CTAAACACTG | ACAAACACGA | GGATTGGGGA | ACGGTGCATC | 960 |
| ACATCAACGA AGCGTGAAAG | CAGAAGCTGC | AGGATGTGAA | GACACGTCAT | GGGGTGGAAT | 1020 |
| TGCGTACTTG GCAGCTTCGT | ATCTCCTTTT | TCTGAGACTG | AACCTGCAGT | CAGGTCCCAC | 1080 |
| AAGGTCAGGT AAAATGGCTC | GATAAAATGT | ACCGTCACTT | TTTGTCGCGT | ATACTGAACT | 1140 |
| CCAAGAGGTC AAAAAAAAA | AAAA | | | | 1165 |

<211> 1165 <212> DNA

<210> 12 <211> 1135 <212> DNA <213> Haematococcus pluvialis

| - | | × |
|------|----------|---|
| 1110 | 00163055 | • |
| wu | 99/63055 | |



| <400> 12 | | | | Jan 12 | |
|--|------------|------------|------------|------------|------|
| CTCGGTAGCT GGCCACAATC | GCTATTTGGA | ACCTGGCCCG | GCGGCAGTCC | GATGCCGCGA | 60 |
| TGCTTCGTTC GTTGCTCAGA | GGCCTCACGC | ATATCCCGCG | CGTGAACTCC | GCCCAGCAGC | 120 |
| CCAGCTGTGC ACACGCGCGA | CTCCAGTTTA | AGCTCAGGAG | CATGCAGCTG | CTTTCCGAGG | 180 |
| ACCGCACAGA CCACATGAGG | GGTGCAAGCA | CCTGGGCAGG | CGGGCAGTCG | CAGGATGAGC | 240 |
| TGATGCTGAA GGACGAGTGC | ATCTTGGTAG | ATGTTGAGGA | CAACATCACA | GGCCATGCCA | 300 |
| GCAAGCTGGA GTGTCACAAG | TTCCTACCAC | ATCAGCCTGC | AGGCCTGCTG | CACCGGGCCT | 360 |
| TCTCTGTGTT CCTGTTTGAC | GATCAGGGGC | GACTGCTGCT | GCAACAGCGT | GCACGCTCAA | 420 |
| AAATCACCTT CCCAAGTGTG | TGGACGAACA | CCTGCTGCAG | CCACCCTTTA | CATGGGCAGA | 480 |
| CCCCAGATGA GGTGGACCAA | CTAAGCCAGG | TGGCCGACGG | AACAGTACCT | GGCGCAAAGG | 540 |
| CTGCTGCCAT CCGCAAGTTG | GAGCACGAGC | TGGGGATACC | AGCGCACCAG | CTGCCGGCAA | 600 |
| GCGCGTTTCG CTTCCTCACG | CGTTTGCACT | ACTGTGCCGC | GGACGTGCAG | CCAGCTGCGA | 660 |
| CACAATCAGC GCTCTGGGGC | GAGCACGAAA | TGGACTACAT | CTTGTTCATC | CGGGCCAACG | 720 |
| TCACCTTGGC GCCCAACCCT | GACGAGGTGG | ACGAAGTCAG | GTACGTGACG | CAAGAGGAGC | 780 |
| TGCGGCAGAT GATGCAGCCG | GACAACGGGC | TTCAATGGTC | GCCGTGGTTT | CGCATCATCG | 840 |
| CCGCGCGCTT CCTTGAGCGT | TGGTGGGCTG | ACCTGGACGC | GGCCCTAAAC | ACTGACAAAC | 900 |
| ACGAGGATTG GGGAACGGTG | CATCACATCA | ACGAAGCGTG | AAGGCAGAAG | CTGCAGGATG | 960 |
| TGAAGACACG TCATGGGGTG | GAATTGCGTA | CTTGGCAGCT | TCGTATCTCC | TTTTTCTGAG | 1020 |
| ACTGAACCTG CAGAGCTAGA | GTCAATGGTG | CATCATATTC | ATCGTCTCTC | TTTTGTTTTA | 1080 |
| GACTAATCTG TAGCTAGAGT | CACTGATGAA | TCCTTTACAA | CTTTCAAAAA | AAAAA ' | 1135 |
| <210> 13 <211> 960 <212> DNA <213> Tagetes erecta <400> 13 | | | | · | |
| CCAAAAACAA CTCAAATCTC | CTCCGTCGCT | CTTACTCCGC | CATGGGTGAC | GACTCCGGCA | 60 |
| TGGATGCTGT TCAGCGACGT | CTCATGTTTG | ACGATGAATG | CATTTTGGTG | GATGAGTGTG | 120 |
| ACAATGTGGT GGGACATGAT | ACCAAATACA | ATTGTCACTT | GATGGAGAAG | ATTGAAACAG | 180 |
| GTAAAATGCT GCACAGAGCA | TTCAGCGTTT | TTCTATTCAA | ТТСААААТАС | GAGTTACTTC | 240 |
| TTCAGCAACG GTCTGCAACC | AAGGTGACAT | TTCCTTTAGT | ATGGACCAAC | ACCTGTTGCA | 300 |
| GCCATCCACT CTACAGAGAA | TCCGAGCTTG | TTCCCGAAAC | GCCTGAGAGA | ATGCTGCACA | 360 |
| GAGGANNNNN NNNNNNNNNN | иииииииии | иииииииии | иииииииии | NNNNNNNNN | 420 |
| инининини инининини | ииииииииии | ииииииииии | иииииииии | иииииииии | 480 |

PCT/US99/12121

WO 99/63055

| иииииииии | иииииииии | ииииииииии | иииииииии | иииииииии | иииииииии | 540 |
|------------|------------|------------|------------|------------|------------|-----|
| ииииииииии | иииииииии | имимимими | ииийииииии | иииииииии | иииииииии | 600 |
| ииииииииии | иииииииии | ииииииииии | NNNNNNNNN | иииииииии | иииииииии | 660 |
| ииииииииии | ииииииииии | TCATGTGCAA | AAGGGTACAC | TCACTGAATG | CAATTTGATA | 720 |
| TGAAAACCAT | ACACAAGCTG | ATATAGAAAC | ACACCCTCAA | CCGAAAAGCA | AGCCTAATAA | 780 |
| TTCGGGTTGG | GTCGGGTCTA | CCATCAATTG | TTTTTTTTTT | TTAACAACTT | TTAATCTCTA | 840 |
| TTTGAGCATG | TTGATTCTTG | TCTTTTGTGT | GTAAGATTTT | GGGTTTCGTT | TCAGTTGTAA | 900 |
| TAATGAACCA | TTGATGGTTT | GCAATTTCAA | GTTCCTATCG | ACATGTAGTG | АТСТААААА | 960 |
| | | | | | | |

<210> 14

<211> 305

<212> PRT

<213> Haematococcus pluvialis

<400> 14

Met Leu Arg Ser Leu Leu Arg Gly Leu Thr His Ile Pro Arg Val Asn 1 5 10 15

Ser Ala Gln Gln Pro Ser Cys Ala His Ala Arg Leu Gln Phe Lys Leu 20 25 30

Arg Ser Met Gln Met Thr Leu Met Gln Pro Ser Ile Ser Ala Asn Leu 35 40 45

Ser Arg Ala Glu Asp Arg Thr Asp His Met Arg Gly Ala Ser Thr Trp 50 55 60

Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys Asp Glu Cys Ile 65 70 75 80

Leu Val Asp Val Glu Asp Asn Ile Thr Gly His Ala Ser Lys Leu Glu 85 90 95

Cys His Lys Phe Leu Pro His Gln Pro Ala Gly Leu Leu His Arg Ala 100 105 110

Phe Ser Val Phe Leu Phe Asp Asp Gln Gly Arg Leu Leu Gln Gln 115 120 125

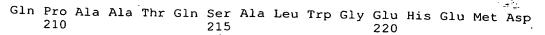
Arg Ala Arg Ser Lys Ile Thr Phe Pro Ser Val Trp Thr Asn Thr Cys 130 135 140

Cys Ser His Pro Leu His Gly Gln Thr Pro Asp Glu Val Asp Gln Leu 145 150 155 160

Ser Gln Val Ala Asp Gly Thr Val Pro Gly Ala Lys Ala Ala Ile 165 170 175

Arg Lys Leu Glu His Glu Leu Gly Ile Pro Ala His Gln Leu Pro Ala 180 185 190

Ser Ala Phe Arg Phe Leu Thr Arg Leu His Tyr Cys Ala Ala Asp Val 195 200 205



- Tyr Ile Leu Phe Ile Arg Ala Asn Val Thr Leu Ala Pro Asn Pro Asp 225 230 235 240
- Glu Val Asp Glu Val Arg Tyr Val Thr Gln Glu Glu Leu Arg Gln Met 245 250 255
- Met Gln Pro Asp Asn Gly Leu Gln Trp Ser Pro Trp Phe Arg Ile Ile 260 265 270
- Ala Ala Arg Phe Leu Glu Arg Trp Trp Ala Asp Leu Asp Ala Ala Leu 275 280 285
- Asn Thr Asp Lys His Glu Asp Trp Gly Thr Val His His Ile Asn Glu 290 295 300

Ala 305

- <210> 15
- <211> 293
- <212> PRT
- <213> Haematococcus pluvialis
- <400> 15
- Met Leu Arg Ser Leu Leu Arg Gly Leu Thr His Ile Pro Arg Val Asn
 1 5 10 15
- Ser Ala Gln Gln Pro Ser Cys Ala His Ala Arg Leu Gln Phe Lys Leu 20 25 30
- Arg Ser Met Gln Leu Ser Glu Asp Arg Thr Asp His Met Arg Gly
 35 40
- Ala Ser Thr Trp Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys 50 55 60
- Asp Glu Cys Ile Leu Val Asp Val Glu Asp Asn Ile Thr Gly His Ala 65 70 75 80
- Ser Lys Leu Glu Cys His Lys Phe Leu Pro His Gln Pro Ala Gly Leu 85 90 95
- Leu His Arg Ala Phe Ser Val Phe Leu Phe Asp Asp Gln Gly Arg Leu 100 105 110
- Leu Leu Gln Gln Arg Ala Arg Ser Lys Ile Thr Phe Pro Ser Val Trp 115 120 125
- Thr Asn Thr Cys Cys Ser His Pro Leu His Gly Gln Thr Pro Asp Glu 130 135 140
- Val Asp Gln Leu Ser Gln Val Ala Asp Gly Thr Val Pro Gly Ala Lys
 145 150 155 160
- Ala Ala Ile Arg Lys Leu Glu His Glu Leu Gly Ile Pro Ala His 165 170 175
- Gln Leu Pro Ala Ser Ala Phe Arg Phe Leu Thr Arg Leu His Tyr Cys



180 185 1

Ala Ala Asp Val Gln Pro Ala Ala Thr Gln Ser Ala Leu Trp Gly Glu 195 200 205

His Glu Met Asp Tyr Ile Leu Phe Ile Arg Ala Asn Val Thr Leu Ala 210 215 220

Pro Asn Pro Asp Glu Val Asp Glu Val Arg Tyr Val Thr Gln Glu Glu 225 230 235 240

Leu Arg Gln Met Met Gln Pro Asp Asn Gly Leu Gln Trp Ser Pro Trp 245 250 255

Phe Arg Ile Ile Ala Ala Arg Phe Leu Glu Arg Trp Trp Ala Asp Leu 260 265 270

Asp Ala Ala Leu Asn Thr Asp Lys His Glu Asp Trp Gly Thr Val His
275 280 285

His Ile Asn Glu Ala 290

<210> 16

<211> 284

<212> PRT

<213> Arabidopsis thaliana

<400> 16

BNSDOCID: <WO 996305541 I

Met Ser Val Ser Ser Leu Phe Asn Leu Pro Leu Ile Arg Leu Arg Ser 1 5 10 15

Leu Ala Leu Ser Ser Ser Phe Ser Ser Phe Arg Phe Ala His Arg Pro 20 25 30

Leu Ser Ser Ile Ser Pro Arg Lys Leu Pro Asn Phe Arg Ala Phe Ser 35 40 45

Gly Thr Ala Met Thr Asp Thr Lys Asp Ala Gly Met Asp Ala Val Gln
50 55 60

Arg Arg Leu Met Phe Glu Asp Glu Cys Ile Leu Val Asp Glu Thr Asp 65 70 75 80

Arg Val Val Gly His Val Ser Lys Tyr Asn Cys His Leu Met Glu Asn 85 90 95

Ile Glu Ala Lys Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe 100 105 110

Asn Ser Lys Tyr Glu Leu Leu Gln Gln Arg Ser Asn Thr Lys Val 115 120 125

Thr Phe Pro Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr 130 135 140

Arg Glu Ser Glu Leu Ile Gln Asp Asn Ala Leu Gly Val Arg Asn Ala 145 150 155 160

Ala Gln Arg Lys Leu Leu Asp Glu Leu Gly Ile Val Ala Glu Asp Val 165 170 175



Pro Val Asp Glu Phe Thr Pro Leu Gly Arg Met Leu Tyr Lys Ala Pro 180 185 190

Ser Asp Gly Lys Trp Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Ile 195 200 205

Val Arg Asp Val Lys Val Gln Pro Asn Pro Asp Glu Val Ala Glu Ile 210 215 220

Lys Tyr Val Ser Arg Glu Glu Leu Lys Glu Leu Val Lys Lys Ala Asp 225 230 235 240

Ala Gly Glu Gly Leu Lys Leu Ser Pro Trp Phe Arg Leu Val Val 245 250 255

Asp Asn Phe Leu Met Lys Trp Trp Asp His Val Glu Lys Gly Thr Leu 260 265 270

Val Glu Ala Ile Asp Met Lys Thr Ile His Lys Leu 275 280

<210> 17

<211> 287

<212> PRT

<213> Clarkia breweri

<400> 17

Met Ser Ser Ser Met Leu Asn Phe Thr Ala Ser Arg Ile Val Ser Leu 1 5 10 15

Pro Leu Leu Ser Ser Pro Pro Ser Arg Val His Leu Pro Leu Cys Phe 20 25 30

Phe Ser Pro Ile Ser Leu Thr Gln Arg Phe Ser Ala Lys Leu Thr Phe 35 40 45

Ser Ser Gln Ala Thr Thr Met Gly Glu Val Val Asp Ala Gly Met Asp 50 55 60

Ala Val Gln Arg Arg Leu Met Phe Glu Asp Glu Cys Ile Leu Val Asp 65 70 75 80

Glu Asn Asp Lys Val Val Gly His Glu Ser Lys Tyr Asn Cys His Leu 85 90 95

Met Glu Lys Ile Glu Ser Glu Asn Leu Leu His Arg Ala Phe Ser Val 100 105 110

Phe Leu Phe Asn Ser Lys Tyr Glu Leu Leu Gln Gln Arg Ser Ala 115 120 125

Thr Lys Val Thr Phe Pro Leu Val Trp Thr Asn Thr Cys Cys Ser His 130 135 140

Pro Leu Tyr Arg Glu Ser Glu Leu Ile Asp Glu Asn Cys Leu Gly Val 145 150 155 160

Arg Asn Ala Ala Gln Arg Lys Leu Leu Asp Glu Leu Gly Ile Pro Ala 165 170 175

Glu Asp Leu Pro Val Asp Gln Phe Ile Pro Leu Ser Arg Ile Leu Tyr

180 185 190

Lys Ala Pro Ser Asp Gly Lys Trp Gly Glu His Glu Leu Asp Tyr Leu 195 200 205

Leu Phe Ile Ile Arg Asp Val Asn Leu Asp Pro Asn Pro Asp Glu Val 210 215 220

Ala Glu Val Lys Tyr Met Asn Arg Asp Asp Leu Lys Glu Leu Leu Arg 225 230 235 240

Lys Ala Asp Ala Glu Glu Glu Gly Val Lys Leu Ser Pro Trp Phe Arg 245 250 255

Leu Val Val Asp Asn Phe Leu Phe Lys Trp Trp Asp His Val Glu Lys 260 265 270

Gly Ser Leu Lys Asp Ala Ala Asp Met Lys Thr Ile His Lys Leu 275 280 285

<210> 18

<211> 261

<212> PRT

<213> Arabidopsis thaliana

<400> 18

Thr Gly Pro Pro Pro Arg Phe Phe Pro Ile Arg Ser Pro Val Pro Arg
1 5 10 15

Thr Gln Leu Phe Val Arg Ala Phe Ser Ala Val Thr Met Thr Asp Ser 20 25 30

Asn Asp Ala Gly Met Asp Ala Val Gln Arg Arg Leu Met Phe Glu Asp 35 40 45

Glu Cys Ile Leu Val Asp Glu Asn Asn Arg Val Val Gly His Asp Thr
50 55 60

Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala Glu Asn Leu Leu 65 70 75 80

His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu Leu Leu 85 90 95

Leu Gln Gln Arg Ser Lys Thr Lys Val Thr Phe Pro Leu Val Trp Thr
100 105 110

Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser Glu Leu Ile Glu 115 120 125

Glu Asn Val Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu Phe Asp 130 135 140

Glu Leu Gly Ile Val Ala Glu Asp Val Pro Val Asp Glu Phe Thr Pro 145 150 155 160

Leu Gly Arg Met Leu Tyr Lys Ala Pro Ser Asp Gly Lys Trp Gly Glu
165 170 175

His Glu Val Asp Tyr Leu Leu Phe Ile Val Arg Asp Val Lys Leu Gln 180 185 190 Pro Asn Pro Asp Glu Val Ala Glu Ile Lys Tyr Val Ser Arg Glu Glu 195 200 205

Leu Lys Glu Leu Val Lys Lys Ala Asp Ala Gly Asp Glu Ala Val Lys 210 220

Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe Leu Met Lys Trp 225 230 235 240

Trp Asp His Val Glu Lys Gly Thr Ile Thr Glu Ala Ala Asp Met Lys 245 250 255

Thr Ile His Lys Leu 260

<210> 19

<211> 288

<212> PRT

<213> Saccharomyces cerevisiae

<400> 19

Met Thr Ala Asp Asn Asn Ser Met Pro His Gly Ala Val Ser Ser Tyr 1 5 10 15

Ala Lys Leu Val Gl
n Asn Gl
n Thr Pro Glu Asp Ile Leu Glu Glu Phe 20
 25
 30

Pro Glu Ile Ile Pro Leu Gln Gln Arg Pro Asn Thr Arg Ser Ser Glu 35 40 45

Thr Ser Asn Asp Glu Ser Gly Glu Thr Cys Phe Ser Gly His Asp Glu
50 55 60

Glu Gln Ile Lys Leu Met Asn Glu Asn Cys Ile Val Leu Asp Trp Asp 65 70 75 80

Asp Asn Ala Ile Gly Ala Gly Thr Lys Lys Val Cys His Leu Met Glu 85 90 95

Asn Ile Glu Lys Gly Leu Leu His Arg Ala Phe Ser Val Phe Ile Phe 100 105 110

Asn Glu Gln Gly Glu Leu Leu Gln Gln Arg Ala Thr Glu Lys Ile 115 120 125

Thr Phe Pro Asp Leu Trp Thr Asn Thr Cys Cys Ser His Pro Leu Cys 130 135 140

Ile Asp Asp Glu Leu Gly Leu Lys Gly Lys Leu Asp Asp Lys Ile Lys 145 150 155 160

Gly Ala Ile Thr Ala Ala Val Arg Lys Leu Asp His Glu Leu Gly Ile 165 170 175

Pro Glu Asp Glu Thr Lys Thr Arg Gly Lys Phe His Phe Leu Asn Arg 180 185 190

Ile His Tyr Met Ala Pro Ser Asn Glu Pro Trp Gly Glu His Glu Ile 195 200 205

Asp Tyr Ile Leu Phe Tyr Lys Ile Asn Ala Lys Glu Asn Leu Thr Val

/63055

215 220

Asn Pro Asn Val Asn Glu Val Arg Asp Phe Lys Trp Val Ser Pro Asn 225 230 235 240

Asp Leu Lys Thr Met Phe Ala Asp Pro Ser Tyr Lys Phe Thr Pro Trp 245 250 255

Phe Lys Ile Ile Cys Glu Asn Tyr Leu Phe Asn Trp Trp Glu Gln Leu 260 (265 270

Asp Asp Leu Ser Glu Val Glu Asn Asp Arg Gln Ile His Arg Met Leu 275 280 285

<210> 20

<211> 456

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Consensus
 sequence of four plant B-cyclases

<400> 20

Met Asp Thr Leu Leu Lys Thr Pro Asn Leu Glu Phe Leu Pro His Gly
1 5 10 15

Phe Val Lys Ser Phe Ser Lys Phe Gly Lys Cys Glu Gly Val Cys Val 20 25 30

Lys Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr Lys Lys Glu Asn $35 \hspace{1cm} 40 \hspace{1cm} 45$

Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys Gly Val Val Asp
50 55 60

Leu Ala Val Val Gly Gly Pro Ala Gly Leu Ala Val Ala Gln Gln 65 70 75 80.

Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile Asp Pro Pro Lys Leu 85 90 95

Ile Trp Pro Asn Asn Tyr Gly Val Trp Val Asp Glu Phe Glu Ala Met 100 105 110

Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp Ser Gly Ala Val Tyr Ile 115 120 125

Asp Asp Thr Lys Asp Leu Arg Pro Tyr Gly Arg Val Asn Arg Lys Gln 130 135 140

Leu Lys Ser Lys Met Met Gln Lys Cys Ile Asn Gly Val Lys Phe His 145 150 155 160

Gln Ala Lys Val Ile Lys Val Ile His Glu Glu Lys Ser Met Leu Ile 165 170 175

Cys Asn Asp Gly Thr Ile Gln Ala Thr Val Val Leu Asp Ala Thr Gly 180 185 190

Phe Ser Arg Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln

195 200 205

Val Ala Tyr Gly Ile Leu Ala Glu Val Glu Glu His Pro Phe Asp Lys 210 215 220

Met Val Phe Met Asp Trp Arg Asp Ser His Leu Asn Asn Glu Leu Lys 225 230 235 240

Glu Arg Asn Ser Ile Pro Thr Phe Leu Tyr Ala Met Pro Phe Ser Ser 245 250 255

Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg Pro Gly Leu 260 265 270

Arg Met Asp Asp Ile Gln Glu Arg Met Val Ala Arg Leu His Leu Gly 275 280 285

Ile Lys Val Lys Ser Ile Glu Glu Asp Glu His Cys Val Ile Pro Met 290 295 300

Gly Gly Pro Leu Pro Val Leu Pro Gln Arg Val Val Gly Ile Gly Gly 305 310 315 320

Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met Val Ala Arg Thr 325 330 335

Leu Ala Ala Pro Val Val Ala Asn Ala Ile Ile Tyr Leu Gly Ser 340 345 350

Glu Ser Ser Gly Glu Leu Ser Ala Glu Val Trp Lys Asp Leu Trp Pro 355 360 365

Ile Glu Arg Arg Gln Arg Glu Phe Phe Cys Phe Gly Met Asp Ile 370 375 380

Leu Leu Lys Leu Asp Leu Pro Ala Thr Arg Arg Phe Phe Asp Ala Phe 385 390 395 400

Phe Asp Leu Glu Pro Arg Tyr Trp His Gly Phe Leu Ser Ser Arg Leu 405 410 415

Phe Leu Pro Glu Leu Ile Val Phe Gly Leu Ser Leu Phe Ser His Ala 420 425 430

Ser Asn Thr Ser Arg Glu Ile Met Thr Lys Gly Thr Pro Leu Val Met 435 440 445

Ile Asn Asn Leu Leu Gln Asp Glu 450 455

<210> 21

<211> 524

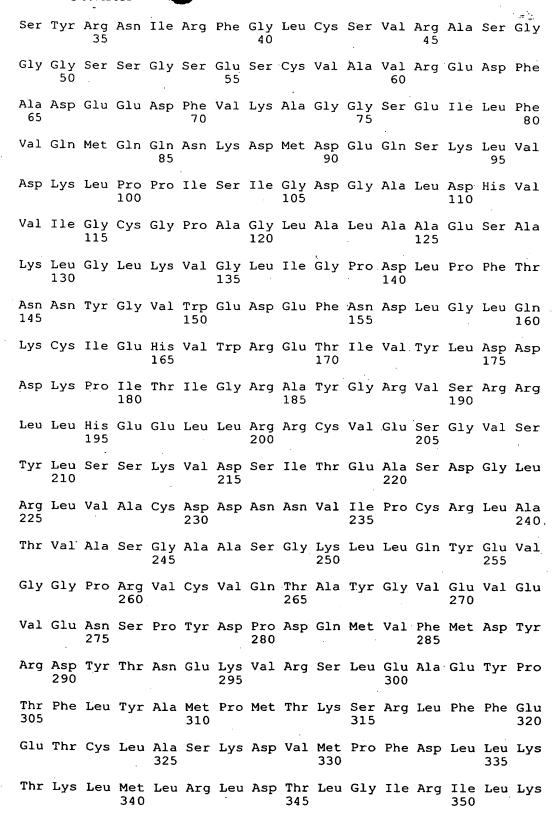
<212> PRT

<213> Arabidopsis thaliana

<400> 21

Met Glu Cys Val Gly Ala Arg Asn Phe Ala Ala Met Ala Val Ser Thr

Phe Pro Ser Trp Ser Cys Arg Arg Lys Phe Pro Val Val Lys Arg Tyr 20 25 30



Thr Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro

WO 99/63055 PCT/US99/12121

| • | | | i |
|--|---------------------------|------------------------|------------|
| 355 | 360 | 365 | æig |
| Asn Thr Glu Gln Lys Asn Lo | eu Ala Phe Gly Ala 75 | Ala Ala Ser Met 380 | Val |
| His Pro Ala Thr Gly Tyr Se 385 390 | er Val Val Arg Ser 395 | Leu Ser Glu Ala | Pro 400 |
| Lys Tyr Ala Ser Val Ile A 405 | la Glu Ile Leu Arg 410 | Glu Glu Thr Thr 415 | Lys |
| Gln Ile Asn Ser Asn Ile Se 420 | er Arg Gln Ala Trp 425 | Asp Thr Leu Trp 430 | Pro · |
| Pro Glu Arg Lys Arg Gln A: 435 | rg Ala Phe Phe Leu 440 | Phe Gly Leu Ala | Leu |
| Ile Val Gln Phe Asp Thr G 450 | lu Gly Ile Arg Ser 55 | Phe Phe Arg Thr 460 | Phe |
| Phe Arg Leu Pro Lys Trp Me 465 470 | et Trp Gln Gly Phe 475 | Leu Gly Ser Thr | Leu 480 |
| Thr Ser Gly Asp Leu Val Le | eu Phe Ala Leu Tyr 490 | Met Phe Val Ile 495 | Ser |
| Pro Asn Asn Leu Arg Lys G. 500 | ly Leu Ile Asn His 505 | Leu Ile Ser Asp 510 | Pro |
| Thr Gly Ala Thr Met Ile Ly 515 | ys Thr Tyr Leu Lys 520 | Val | |
| <210> 22 <211> 1898 <212> DNA <213> Adonis palaestina | | | |
| <400> 22 | | | |
| AAAGGAGTGT TCTATTAATG TTAG | CTGTCGC ATTCTTGCAA | CACTTATATT CAAAC | CTCCAT 60 |
| TTTCTTCTTT TCTCTTCAAA ACAA | ACAAACT AATGTGAGCA | GAGTATCTGG CTATO | GGAACT 120 |
| ACTTGGTGTT CGCAACCTCA TCTC | CTTCTTG CCCTGTGTGG | ACTTTTGGAA CAAGA | AAACCT 180 |
| TAGTAGTTCA AAACTAGCTT ATAA | ACATACA TCGATATGGT | TCTTCTTGTA GAGTA | GATTT 240 |
| TCAAGTGAGA GCTGATGGTG GAAG | GCGGGAG TAGAAGTTCT | GTTGCTTATA AAGAG | GGTTT 300 |
| TGTGGATGAA GAGGATTTTA TCAA | AAGCTGG TGGTTCTGAG | CTTTTGTTTG TCCA | ATGCA 360 |
| GCAAACAAAG TCTATGGAGA AACA | AGGCCAA GCTCGCCGAT | AAGTTGCCAC CAATA | ACCTTT 420 |
| TGGAGAATCC GTGATGGACT TGG | TTGTAAT AGGTTGTGGA | CCTGCTGGTC TTTCF | ACTGGC 480 |
| TGCAGAAGCT GCTAAGCTAG GGT | TGAAAGT TGGCCTTATT | GGTCCTGATC TTCCT | TTTTAC 540 |
| AAATAATTAT GGTGTGTGGG AAGA | ACGAGTT CAAAGATCTT | GGACTTGAAC GTTG1 | TATCGA 600 |

660

720

GCATGCTTGG AAGGACACCA TCGTATATCT TGATAATGAT GCTCCTGTCC TTATTGGTCG

TGCATATGGA CGAGTTAGTC GACATTTGCT ACATGAGGAG TTGCTGAAAA GGTGTGTGGA

BNSDOCID: <WO 996305541 (>

WO 99/63055



| PC | ГЛ | JS99 |)/1 | 21 | 2 |
|----|----|------|-----|----|---|

| GTCAGGTGTA | TCATATCTTG | ATTCTAAAGT | GGAAAGGATC | ACTGAAGCTG | GTGATGGCCA | 780 |
|------------|------------|------------|------------|------------|------------|-------|
| TAGCCTTGTA | GTTTGTGAAA | ATGAGATCTT | TATCCCTTGC | AGGCTTGCTA | CTGTTGCATC | 840 |
| TGGAGCAGCT | TCAGGGAAAC | TTTTGGAGTA | TGAAGTAGGT | GGCCCTCGTG | TTTGTGTCCA | . 900 |
| AACCGCTTAT | GGGGTGGAGG | TTGAGGTGGA | GAACAATCCA | TACGATCCCA | ACTTAATGGT | 960 |
| ATTCATGGAC | TACAGAGACT | ATATGCAACA | GAAATTACAG | TGCTCGGAAG | AAGAATATCC | 1020 |
| AACATTTCTC | TATGTCATGC | CCATGTCGCC | AACAAGACTT | TTTTTTGAGG | AAACCTGTTT | 1080 |
| GGCCTCAAAA | GATGCCATGC | CATTCGATCT | ACTGAAGAGA | AAACTGATGT | CACGATTGAA | 1140 |
| GACTCTGGGT | ATCCAAGTTA | CAAAAGTTTA | TGAAGAGGAA | TGGTCATATA | TTCCTGTTGG | 1200 |
| TGGTTCTTTA | CCAAACACAG | AGCAAAAGAA | CCTAGCATTT | GGTGCTGCAG | CAAGCATGGT | 1260 |
| GCATCCAGCA | ACAGGCTATT | CGGTTGTACG | GTCACTGTCA | GAAGCTCCAA | AATATGCTTC | 1320 |
| TGTAATTGCA | AAGATTTTGA | AGCAAGATAA | CTCTGCGTAT | GTGGTTTCTG | GACAAAGTAG | 1380 |
| TGCAGTAAAC | ATTTCAATGC | AAGCATGGAG | CAGTCTTTGG | CCAAAGGAGC | GAAAACGTCA | 1440 |
| AAGAGCATTC | TTTCTTTTTG | GATTAGAGCT | TATTGTGCAG | CTAGATATTG | AAGCAACCAG | 1500 |
| AACATTCTTT | AGAACCTTCT | TCCGCTTGCC | AACTTGGATG | TGGTGGGGTT | TCCTTGGGTC | 1560 |
| TTCACTATCA | TCTTTCGATC | TCGTCTTGTT | TTCCATGTAC | ATGTTTGTTT | TGGCGCCAAA | 1620 |
| CAGCATGAGG | ATGTCACTTG | TGAGACATTT | GCTTTCAGAT | CCTTCTGGTG | CAGTTATGGT | 1680 |
| AAGAGCTTAC | CTCGAAAGGT | AGTCTCATCT | ATTATTAAAC | TCTAGTGTTT | CACCAAATAA | 1740 |
| ATGAGGATCC | TTCGAATGTG | TATATGATCA | TCTCTATGTA | TATCCTGTAC | TCTAATCTCA | 1800 |
| TAAAGTAAAT | GCCGGGTTTG | ATATTGTTGT | GTCAAACCGG | CCAATGATAT | AAAGTAAATT | 1860 |
| TATTGATACA | AAAGTAGTTT | TTTTCCTTAA | AAAAAAA | | (| 1898 |

<210> 23

<211> 529

<212> PRT

<213> Adonis palaestina

<400> 23

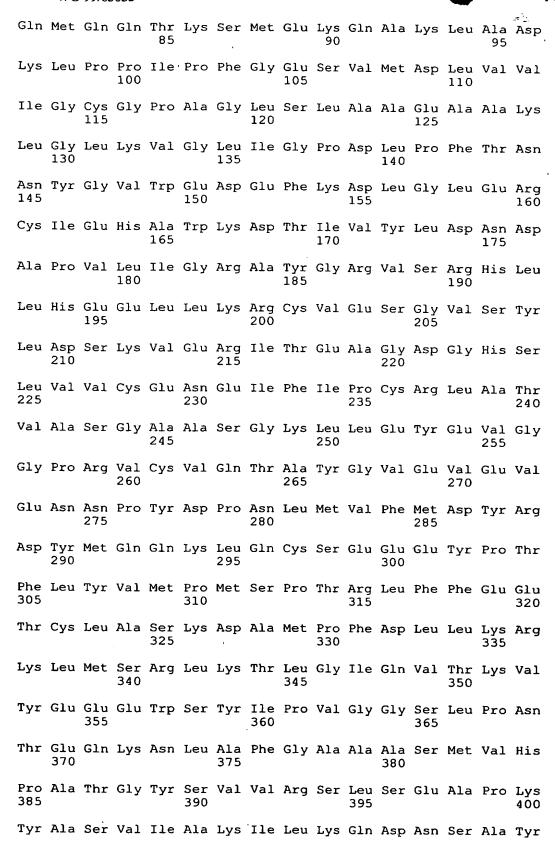
Met Glu Leu Leu Gly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp 1 5 10 15

Thr Phe Gly Thr Arg Asn Leu Ser Ser Ser Lys Leu Ala Tyr Asn Ile 20 25 30

His Arg Tyr Gly Ser Ser Cys Arg Val Asp Phe Gln Val Arg Ala Asp 35 40 45

Gly Gly Ser Gly Ser Arg Ser Ser Val Ala Tyr Lys Glu Gly Phe Val 50 55 60

Asp Glu Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu Leu Leu Phe Val 65 70 75 80





405

410

Val Val Ser Gly Gln Ser Ser Ala Val Asn Ile Ser Met Gln Ala Trp 420 425 430

Ser Ser Leu Trp Pro Lys Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu 435 440 445

Phe Gly Leu Glu Leu Ile Val Gln Leu Asp Ile Glu Ala Thr Arg Thr 450 455 460

Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp Trp Gly Phe 465 470 475 480

Leu Gly Ser Ser Leu Ser Ser Phe Asp Leu Val Leu Phe Ser Met Tyr 485 490 495

Met Phe Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His 500 505 510

Leu Leu Ser Asp Pro Ser Gly Ala Val Met Val Arg Ala Tyr Leu Glu 515 520 525

Arg

DRICHOCID- >WO GODDNEERT I -

<210> 24

<211> 1370

<212> DNA

<213> Potato

<400> 24

TAGCGGAGGA TGAGTTCAAA GATCTTGGTC TTCAAGCCTG CATTGAACAT GTTTGGCTGG 60 GATACCATTG TATATCTTGA TGATGATGAT CCTATTCTTA TTGGCCGTGC CTATGGAAGA 120 GTTAGTCGCC ATTTACTGCA CGAGGAGTTA CTCAAAAGGT GTGTGGAGGC AGGTGTTTTG 180 TATCTAAACT CGAAAGTGGA TAGGATTGTT GAGGCCACAA ATGGCCACAG TCTTGTAGAG 240 TGCGAGGGTG ATGTTGTGAT TCCCTGCAGG TTTGTGACTG TTGCATCGGG AGCAGCCTCG 300 GGGAAATTCT TGCAGTATGA GTTGGGAGGT CCTAGAGTTT CTGTTCAAAC AGCTTATGGA 360 GTGGAAGTTG AGGTCGATAA CAATCCATTT GACCCGAGCC TGATGGTTTT CATGGATTAT 420 AGAGACTATG TCAGACACGA CGCTCAATCT TTAGAAGCTA AATATCCAAC ATTTCTCTAT 480 GCCATGCCCA TGTCTCCAAC ACGAGTCTTT TTCGAGGAAA CTTGTTTGGC TTCAAAAGAT 540 GCAATGCCAT TCGATCTGTT AAAGAAAAA TTGATGTTAC GATTGAACAC CCTCGGTGTA 600 AGAATTAAAG AAATTTATGA GGAGGAATGG TCTTACATAC CAGTTGGAGG ATCTTTGCCA 660 AATACAGAAC AAAAAACACT TGCATTTGGT GCTGCTGCTA GCATGGTTCA TCCAGCCACA 720 GGTTATTCAG TCGTCAGATC ACTGTCTGAA GCTCCAAAAT GCGCCTTCGT GCTTGCAAAT 780 ATATTACGAC AAAATCATAG CAAGAATATG CTTACTAGTT CAAGTACCCC GAGTATTTCA 840 ACTCAAGCTT GGAACACTCT TTGGCCACAA GAACGAAAAC GACAAAGATC GTTTTTCCTA 900

| ' | | |
|-------------|--|----------------|
| WO 99/63055 | | PCT/US99/12121 |

| TTTGGACTGG CTCTGATATT GCAGCTGGAT ATTGAGGGGA TAAGGTCATT TTTCCGC | GCG | | | | | | | | |
|--|-----|--|--|--|--|--|--|--|--|
| TTCTTCCGTG TGCCAAAATG GATGTGGCAG GGATTTCTTG GTTCAAGTCT TTCTTAG | CAG | | | | | | | | |
| ACCTCATGTT ATTTGCCTTC TACATGTTTA TTATTGCACC AAATGACATG AGAAGAG | GCT | | | | | | | | |
| TAATCAGACA TCTTTTATCT GATCCTACTG GTGCAACATT GATAAGAACT TATCTTA | CAT | | | | | | | | |
| TTTAGAGTAA ATTCCTCCTA CAATAGTTGT TGAAAGAGGC CTCATTACTT CAGATTC | ATA | | | | | | | | |
| ACAGAAATCG CGGTCTCTCG AGGCCTTGTA TATAACATTT TCACTAGGTT AATATTG | CTT | | | | | | | | |
| GAATAAGTTG CACAGTTTCA GTTTTTGTAT CTGCTTCTTT TTTGTCCAAG ATCATGT | ATT | | | | | | | | |
| GACCAATTTA TATACATTGC CAGTATATAT AAATTTTATA AAAAAAAAA | | | | | | | | | |
| <210> 25 <211> 377 <212> PRT <213> Potato | | | | | | | | | |
| <400> 25 Asp Glu Pho Luc Asp Lou Clu Lou Cla Ala Gua Til Cu Luc Cla Asp Glu Pho Luc Asp Lou Cla Asp Glu Pho Luc Cla Asp Glu Pho Cla Til Cu Til | | | | | | | | | |
| Asp Glu Phe Lys Asp Leu Gly Leu Gln Ala Cys Ile Glu His Val Tr 1 5 10 15 | Þ | | | | | | | | |
| Arg Asp Thr Ile Val Tyr Leu Asp Asp Asp Pro Ile Leu Ile Gl | У | | | | | | | | |
| Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu His Glu Glu Leu Leu 35 40 45 | ט | | | | | | | | |
| Lys Arg Cys Val Glu Ala Gly Val Leu Tyr Leu Asn Ser Lys Val Asp 50 55 60 | Ò | | | | | | | | |
| Arg Ile Val Glu Ala Thr Asn Gly His Ser Leu Val Glu Cys Glu Gl 65 70 75 80 | | | | | | | | | |
| Asp Val Val Ile Pro Cys Arg Phe Val Thr Val Ala Ser Gly Ala Ala 85 90 95 | a ' | | | | | | | | |
| Ser Gly Lys Phe Leu Gln Tyr Glu Leu Gly Gly Pro Arg Val Ser Val 100 105 110 | l | | | | | | | | |
| Gln Thr Ala Tyr Gly Val Glu Val Glu Val Asp Asn Asn Pro Phe Asp 115 120 125 | Þ | | | | | | | | |
| Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp Tyr Val Arg His Asp 130 | Þ | | | | | | | | |
| Ala Gln Ser Leu Glu Ala Lys Tyr Pro Thr Phe Leu Tyr Ala Met Pro 145 150 155 166 | | | | | | | | | |
| Met Ser Pro Thr Arg Val Phe Phe Glu Glu Thr Cys Leu Ala Ser Lys 165 170 175 | 5 | | | | | | | | |
| Asp Ala Met Pro Phe Asp Leu Leu Lys Lys Lys Leu Met Leu Arg Leu 180 185 190 | 1 | | | | | | | | |
| | | | | | | | | | |

Asn Thr Leu Gly Val Arg Ile Lys Glu Ile Tyr Glu Glu Glu Trp Ser 195 200 205

RNSOCIO -WO GORGOREAT I -

WO 99/63055



- Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Thr Leu 210 220
- Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser 225 230 235 240
- Val Val Arg Ser Leu Ser Glu Ala Pro Lys Cys Ala Phe Val Leu Ala 245 250 255
- Asn Ile Leu Arg Gln Asn His Ser Lys Asn Met Leu Thr Ser Ser Ser 260 265 270
- Thr Pro Ser Ile Ser Thr Gln Ala Trp Asn Thr Leu Trp Pro Gln Glu 275 280 285
- Arg Lys Arg Gln Arg Ser Phe Phe Leu Phe Gly Leu Ala Leu Ile Leu 290 295 300
- Gln Leu Asp Ile Glu Gly Ile Arg Ser Phe Phe Arg Ala Phe Phe Arg 305 310 315 320
- Val Pro Lys Met Met Trp Gly Phe Leu Gly Ser Ser Leu Ser Xaa Ala 325 330 335
- Asp Leu Met Leu Phe Ala Phe Tyr Met Phe Ile Ile Ala Pro Asn Asp $340 \hspace{1.5cm} 345 \hspace{1.5cm} 350$
- Met Arg Arg Gly Leu Ile Arg His Leu Leu Ser Asp Pro Thr Gly Ala 355 360 365
- Thr Leu Ile Arg Thr Tyr Leu Thr Phe 370 375
- <210> 26
- <211> 533
- <212> PRT
- <213> Chimeric lettuce/potato
- <400> 26
- Met Glu Cys Phe Gly Ala Arg Asn Met Thr Ala Thr Met Ala Val Phe 1 5 10 15
- Thr Cys Pro Arg Phe Thr Asp Cys Asn Ile Arg His Lys Phe Ser Leu 20 25 30
- Leu Lys Gly Arg Arg Phe Thr Asn Leu Ser Ala Ser Ser Ser Leu Arg 35 40 45
- Gln Ile Lys Cys Ser Ala Lys Ser Asp Arg Cys Val Val Asp Lys Gln 50 55 60
- Gly Ile Ser Val Ala Asp Glu Glu Asp Tyr Val Lys Ala Gly Gly Ser 65 70 75 80
- Glu Leu Phe Phe Val Gln Met Gln Arg Thr Lys Ser Met Glu Ser Gln 85 90 95
- Ser Lys Leu Ser Glu Lys Leu Ala Gln Ile Pro Ile Gly Asn Cys Ile 100 105 110
- Leu Asp Leu Val Val Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala

Ala Glu Ser Ala Lys Leu Gly Leu Asn Val Gly Leu Ile Gly Pro Asp 130 135 140

Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Gln Asp Glu Phe Ile Gly 145 150 155 160

Leu Gly Leu Glu Gly Cys Ile Glu His Ser Trp Lys Asp Thr Leu Val 165 170 175

Tyr Leu Asp Asp Ala Asp Pro Ile Arg Ile Gly Arg Ala Tyr Gly Arg 180 185 190

Val His Arg Asp Leu Leu His Glu Glu Leu Leu Arg Arg Cys Val Glu 195 200 205

Ser Gly Val Ser Tyr Leu Ser Ser Lys Val Glu Arg Ile Thr Glu Ala 210 215 220

Pro Asn Gly Tyr Ser Leu Ile Glu Cys Glu Gly Asn Ile Thr Ile Pro 225 230 235 240

Cys Arg Leu Ala Thr Val Ala Ser Gly Ala Ala Ser Gly Lys Phe Leu 245 250 255

Glu Tyr Glu Leu Gly Gly Pro Arg Val Ser Val Gln Thr Ala Tyr Gly 260 265 270

Val Glu Val Glu Val Asp Asn Asn Pro Phe Asp Pro Ser Leu Met Val 275 280 285

Phe Met Asp Tyr Arg Asp Tyr Val Arg His Asp Ala Gln Ser Leu Glu 290 295 300

Ala Lys Tyr Pro Thr Phe Leu Tyr Ala Met Pro Met Ser Pro Thr Arg 305 310 315 320

Val Phe Phe Glu Glu Thr Cys Leu Ala Ser Lys Asp Ala Met Pro Phe 325 330 335

Asp Leu Leu Lys Lys Lys Leu Met Leu Arg Leu Asn Thr Leu Gly Val 340 345 350

Arg Ile Lys Glu Ile Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly 355 360 365

Gly Ser Leu Pro Asn Thr Glu Gln Lys Thr Leu Ala Phe Gly Ala Ala 370 375 380

Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu 385 390 395 400

Ser Glu Ala Pro Lys Cys Ala Phe Val Leu Ala Asn Ile Leu Arg Gln 405 410 415

Asn His Ser Lys Asn Met Leu Thr Ser Ser Ser Thr Pro Ser Ile Ser 420 425 430

Thr Gln Ala Trp Asn Thr Leu Trp Pro Gln Glu Arg Lys Arg Gln Arg 435 440 445

WO 99/63055

Ser Phe Phe Leu Phe Gly Leu Ala Leu Ile Leu Gln Leu Asp Ile Glu 450 455 460

Gly Ile Arg Ser Phe Phe Arg Ala Phe Phe Arg Val Pro Lys Trp Met 465 470 475 480

Trp Gln Gly Phe Leu Gly Ser Ser Leu Ser Xaa Ala Asp Leu Met Leu 485 490 495

Phe Ala Phe Tyr Met Phe Ile Ile Ala Pro Asn Asp Met Arg Arg Gly 500 505 510

Leu Ile Arg His Leu Leu Ser Asp Pro Thr Gly Ala Thr Leu Ile Arg 515 520 525

Thr Tyr Leu Thr Phe 530

<210> 27

<211> 374

<212> PRT

<213> Arabidopsis thaliana

<400> 27

Glu Asp Glu Phe Asn Asp Leu Gly Leu Gln Lys Cys Ile Glu His Val
1 5 10 15

Trp Arg Glu Thr Ile Val Tyr Leu Asp Asp Lys Pro Ile Thr Ile
20 25 30

Gly Arg Ala Tyr Gly Arg Val Ser Arg Arg Leu Leu His Glu Glu Leu 35 40 45

Leu Arg Arg Cys Val Glu Ser Gly Val Ser Tyr Leu Ser Ser Lys Val
50 60

Asp Ser Ile Thr Glu Ala Ser Asp Gly Leu Arg Leu Val Ala Cys Asp 65 70 75 80

Asp Asn Asn Val Ile Pro Cys Arg Leu Ala Thr Val Ala Ser Gly Ala 85 90 95

Ala Ser Gly Lys Leu Leu Gln Tyr Glu Val Gly Gly Pro Arg Val Cys 100 105 110

Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Glu Asn Ser Pro Tyr 115 120 125

Asp Pro Asp Gln Met Val Phe Met Asp Tyr Arg Asp Tyr Thr Asn Glu 130 135 140

Lys Val Arg Ser Leu Glu Ala Glu Tyr Pro Thr Phe Leu Tyr Ala Met 145 150 155 160

Pro Met Thr Lys Ser Arg Leu Phe Phe Glu Glu Thr Cys Leu Ala Ser 165 170 175

Lys Asp Val Met Pro Phe Asp Leu Leu Lys Thr Lys Leu Met Leu Arg 180 185 190

Leu Asp Thr Leu Gly Ile Arg Ile Leu Lys Thr Tyr Glu Glu Glu Trp

540

600

195 200 205 Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Asn 210 215 Leu Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr 235 Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys Tyr Ala Ser Val Ile 245 255 Ala Glu Ile Leu Arg Glu Glu Thr Thr Lys Gln Ile Asn Ser Asn Ile 260 Ser Arg Gln Ala Trp Asp Thr Leu Trp Pro Pro Glu Arg Lys Arg Gln 280 Arg Ala Phe Phe Leu Phe Gly Leu Ala Leu Ile Val Gln Phe Asp Thr Glu Gly Ile Arg Ser Phe Phe Arg Thr Phe Phe Arg Leu Pro Lys Trp 310 320 Met Trp Gln Gly Phe Leu Gly Ser Thr Leu Thr Ser Gly Asp Leu Val 325 Leu Phe Ala Leu Tyr Met Phe Val Ile Ser Pro Asn Asn Leu Arg Lys 345 Gly Leu Ile Asn His Leu Ile Ser Asp Pro Thr Gly Ala Thr Met Ile 355 Lys Thr Tyr Leu Lys Val 370 <210> 28 <211> 1002 <212> DNA <213> Adonis palaestina <400> 28 ATTCATCTTC AGCAGCGCTG TCGTACTCTT TCTATATCTT CTTCCATCAC TAACAGTAGT 60 CGCCGACGGT TGAATCGGCT ATTCGCCTCA ACGTCAACTA TGGGTGAAGT CACTGATGCT 120 GGAATGGATG CTGTTCAGAA GCGGCTCATG TTCGACGACG AATGTATTTT GGTGGATGAG 180 AATGACAAGG TCGTCGGGCA TGATTCCAAA TACAACTGTC ATTTGATGGA AAAGATAGAG 240 GCAGAAAATT TGCTTCACAG AGCCTTCAGT GTTTTCTTGT TCAACTCAAA ATATGAATTG 300 CTTCTTCAGC AACGATCCGC CACAAAGGTA ACATTCCCGC TCGTATGGAC AAACACATGT 360 TGCAGTCATC CTCTCTTTCG TGATTCCGAG CTCATAGAAG AAAATTATCT CGGTGTACGA 420 AACGCTGCAC AAAGAAAGCT TTTAGACGAG CTAGGCATTC CAGCTGAAGA TGTCCCAGTT

GATGAATTTA CTCCTCTTGG TCGCATTCTT TACAAAGCTC CATCTGACGG CAAATGGGGA

GAGCACGAAT TGGACTATCT CCTATTTATT GTCCGAGATG TGAAATACGA TCCAAACCCA

WO 99/63055 PCT/US99/12121

| | | | | | ** | 170377/121 |
|---|-------------------|------------|--------------|------------|--------------|------------|
| GATGAAGTT | G CTGATGCTAA | GTATGTTAA | r cgcgaggagi | TGAGAGAGA' | I ACTGAGAAAA | 660 |
| GCTGATGCT | G GTGAAGAGGG | ACTCAAGTTO | TCTCCTTGGT | TTAGATTGG | r tgttgataac | 720 |
| TTTTTGTTC | A AGTGGTGGGA | TCATGTAGAG | CAGGGTACGA | TTAAGGAAG | T TGCTGACATG | 780 |
| AAAACTATC | C ACAAGTTGAC | TTAAGAGGAC | TTCTCTCCTC | TGTTCTACTA | A TTTGTTTTT | 840 |
| GCTACAATA | A GTGGGTGGTG | ATAAGCAGTI | TTTCTGTTTT | CTTTAATTT | A TGGCTTTTGA | 900 |
| ATTTGCCTC | G ATGTTGAACT | TGTAACATAT | TTAGACAAAT | ATGAGACCTI | GTAAGTTGAA | 960 |
| TTTGAGGCT | AATTTATATT | TTTGGGAACA | TAATAATGTT | AA | | 1002 |
| <210> 29 <211> 1271 <212> DNA <213> Ador | l nis palaesti | na | | | | . * |
| <400> 29 | | • | | • | | |
| | | | | | TACAAAAGTT | 60 |
| | | | | | GATACGCATC | 120 |
| | | | | | CGGAATCCCG | 180 |
| And the second second second | | | | | CACCACTAAA | 240 |
| | CTTCGTGTTC | | | | | 300 |
| TTCTATTTCT | TCTTCCATCA | CTAACAGTCC | TCGCCGAGGG | TTGAATCGGC | TGTTCGCCTC | 360 |
| AACGTCGACT | ATGGGTGAAG | TCGCTGATGC | TGGTATGGAT | GCCGTCCAGA | AGCGGCTTAT | 420 |
| GTTCGACGAT | GAATGTATTT | TGGTGGATGA | GAATGACAAG | GTCGTCGGAC | ATGATTCCAA | 480 |
| ATACAACTGT | CATTTGATGG | AAAAGATAGA | GGCAGAAAAC | TTGCTTCACA | GAGCCTTCAG | 540 |
| TGTTTTCTTA | TTCAACTCAA | AATACGAGTT | GCTTCTTCAG | CAACGATCTG | CAACGAAGGT | 600 |
| AACATTCCCG | CTCGTATGGA | CAAACACCTG | TTGCAGCCAT | CCCCTCTTCC | GTGATTCCGA | 660 |
| ACTCATAGAA | GAAAATTTTC | TCGGGGTACG | ÄAACGCTGCA | CAAAGGAAGC | TTTTAGACGA | 720 |
| GCTAGGCATT | CCAGCTGAAG | ACGTACCAGT | TGATGAATTC | ACTCCTCTTG | GTCGCATTCT | 780 |
| TTACAAAGCT | CCATCTGACG | GAAAATGGGG | AGAGCACGAA | CTGGACTATC | TTCTGTTTAT | 840 |
| TGTCCGAGAT | GTGAAATACG | ATCCAAACCC | AGATGAAGTT | GCTGACGCTA | AGTACGTTAA | 900 |
| TCGCGAGGAG | TTGAAAGAGA | TACTGAGAAA | AGCTGATGCA | GGTGAAGAGG | GAATAAAGTT | 960 |
| GTCTCCTTGG | TTTAGATTGG | TTGTGGATAA | CTTTTTGTTC | AAGTGGTGGG | ATCATGTAGA | 1020 |
| GGAGGGGAAG | ATTAAGGACG | TCGCCGACAT | GAAAACTATC | CACAAGTTGA | CTTAAGAGAA | 1080 |
| | | | | | TGAGCAGTTT | 1140 |
| | TTTAATTTTG | | | | | 1200 |
| | TGAGACCTTG | | | | | 1260 |
| | | | | - | | 1200 |

3NSDOCID: <WO___9963055A1_I_>

- -ΑΑΑΑΑΑΑΑ Α 1271 <210> 30 <211> 1109 <212> DNA <213> Haematococcus pluvialis <400> 30 TGGAACCTGG CCCGGCGGCA GTCCGATGCC GCGATGCTTC GTTCGTTGCT CAGAGGCCTC 60 ACGCATATCC CGCGCGTGAA CTCCGCCCAG CAGCCCAGCT GTGCACACGC GCGACTCCAG 120 TTTAAGCTCA GGAGCATGCA GCTGCTTGCC GAGGACCGCA CAGACCACAT GAGGGGTGCA 180 AGCACCTGGG CAGGCGGCA GTCGCAGGAT GAGCTGATGC TGAAGGACGA GTGCATCTTA 240 GTGGATGCTG ACGACAACAT CACAGGCCAT GCCAGCAAGC TGGAGTGCCA CAAATTCCTA 300 CCACATCAGC CTGCAGGCCT GCTGCACCGG GCCTTCTCTG TGTTCCTGTT TGACGACCAG 360 GGGCGACTGC TGCTGCAACA GCGTGCACGC TCAAAAATCA CCTTCCCAAG TGTGTGGACG 420 AACACCTGCT GCAGCCACCC TCTACATGGG CAGACCCCAG ATGAGGTGGA CCAACTAAGC 480 CAGGTGGCCG ACGGCACAGT ACCTGGCGCA AAAGCTGCTG CCATCCGCAA GTTGGAGCAC 540 GAGCTGGGGA TACCAGCGCA CCAGCTGCCG GCAAGCGCGT TTCGCTTCCT CACGCGTTTG 600 CACTACTGTG CCGCGGACGT GCAGCCGGCT GCGACACAAT CAGCGCTCTG GGGCGAGCAC 660 GAGATGGACT ACATCTTATT CATCCGGGCC AACGTCACCT TGGCGCCCAA CCCTGACGAG 720 GTGGACGAAG TCAGGTACGT GACGCAAGAG GAGCTGCGGC AGATGATGCA GCCGGACAAC 780 GGGTTGCAAT GGTCGCCGTG GTTTCGCATC ATCGCCGCGC GCTTCCTTGA GCGTTGGTGG 840 GCTGACCTGG ACGCGGCCCT AAACACTGAC AAACACGAGG ATTGGGGAAC GGTGCATCAC 900 ATCAACGAAG CGTGAAGGCA GAAGCTGCAG GATGTGAAGA CACGTCATGG GGTGGAATTG 960 CGTACTTGGC AGCTTCGTAT CTCCTTTTTC TGAGACTGAA CCTGCAGAGC TAGAGTCAAT 1020 GGTGCATCAT ATTCATCGTC TCTCTTTTGT TTTAGACTAA TCTGTAGCTA GAGTCACTGA 1080 TGAATCCTTT ACAACTTTCA AAAAAAAA 1109 <210> 31 <211> 985 <212> DNA <213> Lactuca sativa <400> 31 TGCCAAAATG TTGAAATTTC CCCCTTTTAA AACCATTGCT ACCATGATCT CTTCTCCATA 60 TTCTTCCTTC TTGCTGCCTC GGAAATCTTC TTTCCCTCCA ATGCCGTCTC TCGCAGCCGC 120 TAGTGTTTTC CTCCACCCTC TTTCGTCTGC CGCTATGGGC GATTCCAGCA TGGATGCTGT 180 CCAGCGACGT CTCATGTTCG ATGACGAATG CATTTTGGTG GATGAGAATG ACAAAGTGGT 240 TGGCCATGAT ACTAAATACA ATTGTCATTT GATGGAGAAG ATTGAAAAGG GAAATATGCT 300

PCT/US99/12121

WO 99/63055

| 1110 | 99/630 | 122 |
|------|---------|-----|
| w | 44/n.11 | |



PCT/US99/1212

| | • | 7,00 | | | | | |
|---|------------------------------------|------------|------------|------------|------------|------------|-----|
| | ACACAGAGCA | TTCAGTGTGT | TCTTGTTCAA | CTCGAAATAT | GAATTACTCC | TTCAGCAACG | 360 |
| | TTCTGCAACC | AAGGTGACTT | TCCCTTTGGT | ATGGACAAAC | ACGTGTTGCA | GCCATCCACT | 420 |
| | ATACAGGGAG | AGTGAGCTTA | TTGACGAAAA | CGCCCTTGGG | GTGAGGAATG | CTGCACAGAG | 480 |
| | GAAGCTCCTG | GATGAACTCG | GCATCCCTGG | AGCAGATGTT | CCGGTTGATG | AGTTCACTCC | 540 |
| | ATTGGGTCGC | ATTCTATACA | AGGCCGCATC | GGATGGAAAG | TGGGGAGAAC | ATGAACTTGA | 600 |
| | TTACCTGCTG | TTTATGGTAC | GTGATGTTGG | TTTGGATCCG | AACCCAGATG | AAGTGAAAGA | 660 |
| | TGTAAAATAT | GTGAACCGGG | AAGAGCTGAA | GGAATTGGTA | AGGAAGGCGG | ATGCTGGTGA | 720 |
| | AGAGGGTGTG | AAGCTGTCCC | CGTGGTTCAA | ATTGATTGTC | GATAATTTCT | TGTTTCAGTG | 780 |
| | GTGGGATCGA | CTCCATAAGG | GAACCCTAAC | CGAAGCTATT | GATATGAAAA | CAATCCACAA | 840 |
| | ACTCACATAA | AAACACTACA | CTAGTAGGAG | AGAGGATTAT | ATGAGATATT | TGTTATATGT | 900 |
| | GAAATTGAAA | TTCAGATGAA | TGCTTGTÄTT | TATTTCTATT | TGGACAAACT | TCAACTTCTT | 960 |
| | TTTGCTACCT | TATCAGAAAA | AAAAA | • | | | 985 |
| | <210> 32 <211> 988 <212> DNA | | · | | | | |
| , | <213> Lacti | ca sativa | | • | | | |
| | <400> 32 | AAAATCTCTT | ССАТТААСТС | СТСАЛАТСТС | CACCERECCC | CCMCMMAAMO | 60 |
| | | CACTTTCACC | | • | | | 60 |
| | | ACGTCTCATG | | , | | | 120 |
| | | TGATACCAAA | | | | | 180 |
| | | AGCATTCAGT | | | | | 240 |
| | | AACCAAGGTG | 1 | | | • | 300 |
| | | | | • | | | 360 |
| | | AGAATCGGAG | | | | | 420 |
| | | | | | | GATGAGTTCA | 480 |
| | | | • | | | GAACATGAAG | 540 |
| | | | | | | GATGAGGTGG | 600 |
| | | | | | | GCGGATGCGG | 660 |
| | | • | | | | TTCTTGTTCA | 720 |
| | | | | | | AAAACCATTC | 780 |
| | | | | | | ATAATTTGTG | 840 |
| | | TCGGTCCTTC | | | | | 900 |
| | TATTGTGTAC | TTGTAACGTA | GGCCCTTTGG | TTACGCTTTA | AGAGTTTAAT | AAAGAACCAC | 960 |



| ССТТААТТТА АДАДАДАДА ДДАДАДАД | 988 |
|--|-----------------------|
| <210> 33 <211> 1874 <212> DNA <213> Chlamydomonas reinhardtii | · |
| <400> 33 GGCACGAGCT CGAGTTTGTT TTACCATGAC ATCGGGAATT TGGAAG | GCTTG AACTACCTCA 60 |
| ATTACTCAAG TAACTCGCGG CAACACATTT CGCGCGCCAT CGCTGT | TTTTC TCTGCTCCAG 120 |
| CTACCGAGCA GCATTGCTTT AGATCGCTTT GATGTCATAA ACTCCC | CACTT ATATGAGATC 180 |
| CAGTTTCATC GAGCCCAAGC CCAGAGCGCA ACCTGTCTTA AGCCGC | CGGCA GGGCGTCCAT 240 |
| GCGCCTCGCG CAAAGCCGTG CTCTCGTTGC GCGTGTCAGC TCCGCC | CCTGT GGCCGGGAGC 300 |
| AGGACTTTCA CAGGCTCAAA GCGTTGCGGT GCGAATGGCG AGTTCG | STCAA CCTGGGAAGG 360 |
| CACGGGCCTG AGCCAGGATG ACTTCATGCA GCGGGACGAG TGCTTG | GGTGG TGGACGAGCA 420 |
| GGACCGGCTG CTAGGCACCG CCAACAAGTA CGACTGCCAC CGCTTC | CGAGG CGGCCAAGGG 480 |
| CCAGCCCTGC GGCCGCCTGC ACCGCGCCTT CTCCGTGTTC CTGTTC | CAGCC CCGACGGCCG 540 |
| ACTGCTGCTG CAGCAGCGC CAGCCAGCAA GGTGACGTTC CCGGGT | CGTGT GGACCAACAC 600 |
| CTGCTGCTCG CACCCGCTGG CGGGCCAGGC GCCGGACGAG GTGGAC | CCTGC CGGCGGCGGT 660 |
| AGCCTCGGGC CAGGTGCCGG GCATCAAGGC GGCGGCGGTG CGCAAG | GCTGC AGCACGAGCT 720 |
| GGGGATACCG CCGGAGCAGG TTCCCGCCTC CTCCTTCTCC TTCCTC | CACGC GTCTGCACTA 780 |
| CTGCGCCGCC GACACCGCCA CGCACGGCCC GGCGGCGGAG TGGGGC | CGAGC ACGAGGTGGA 840 |
| CTACGTGCTG TTCGTGCGGC CGCAGCAGCC CGTCAGCCTG CAGCCC | CAACC CAGACGAGGT 900 |
| GGACGCCACG CGCTACGTGA CGCTGCCGGA GCTTCAGTCC ATGATG | GGCGG ACCCCGGCCT 960 |
| CAGCTGGAGC CCCTGGTTCC GCATCCTGGC CACACAGCCC GCCTTC | CCTGC CCGCCTGGTG 1020 |
| GGGCGACCTG AAGCGGCGCT GGCGCCCGGG CGGCAGCCGA CTGTCG | GGACT GGGGCACCAT 1080 |
| CCACCGCGTC ATGTGAAGAA AAAGGGGAAG CAGGGGCGGG AGCGGG | GGGAT GAATGGGAAT 1140 |
| GTGAATGCGA TTGTGATGCG GCGTGGGATG AGGTCTGAAG ACAGGG | GGGAA AATCGGGGGG 1200 |
| CGGGCGTGAG CGTGTGTGTA CGTGAGCGAC AAAGCCGGGA GGCGGA | ACCGC GCGATGGGTA 1260 |
| CATGTGTGTG CGGAGGGTCG GTGGGTCGGT CGGTTGCGCG GCATAG | SCGTG TTGTGTGTGT 1320 |
| GCGGCTGCAG GGGTATGTGG GCACCCGGGC ACGGAGGAGA AGGCAC | CACGC AGGTGGCGCG 1380 |
| GAGGTGTGTC AGGGGCCATG GGCGGGCCTC ACTCCTGGTC GTGCCC | CAGTG GTCTCGTGGG 1440 |
| CAGAGTGGCA GGGGCTGCAC CCATATGAGC GGCGCACTGC CGCGCT | RGGGC TAAGTCCTTA 1500 |
| TCACTTGGTG AGGTGGGCG AGGTGGCTGT GGGCGGCGGG CGCAGT | TGGCA GAAGGACACG 1560 |
| GTGTGTGAGC GGTGGAGCTC TGGCCGTGCC GGCCGTGAGG GGCGGA | ATAGC GATATGACGT 1620 |
| | |

WO 99/63055 PCT/US99/12121

| • | | | | • | | |
|--|------------|------------|------------|------------|------------|------|
| TGTGCTTGGC | CGCTGTAATG | CGGGAGAATG | TGCAGGCCGC | GAGAAGCGGG | CGGTGGCAGG | 1680 |
| AGGCCGCAGG | CTGCAGCACC | CGTTGGGGAG | GTGCCACCTG | CAGGCGCGGC | GCCGGGCGG | 1740 |
| CCTGAGTAAT | GGGCGCCTGA | GTAGTGGCGG | CCACAGGAGG | CGCAGGAGGC | AGCAGCAGGA | 1800 |
| GGACGAGCTG | GAGGGACCCG | TTGGCAACCC | AAGGTTGCGC | GTGTAACATA | GTGGCCATAC | 1860 |
| AAAAAAAA | AAAA | | | | | 1874 |
| <210> 34 <211> 954 <212> DNA <213> Taget | les erecta | | | | | |
| <400> 34 CCAAAAACAA | CTCAAATCTC | CTCCGTCGCT | CTTACTCCGC | CATGGGTGAC | GACTCCGGCA | 60 |
| TGGATGCTGT | TCAGCGACGT | CTCATGTTTG | ACGATGAATG | CATTTTGGTG | GATGAGTGTG | 120 |
| ACAATGTGGT | GGGACATGAT | ACCAAATACA | ATTGTCACTT | GATGGAGAAG | ATTGAAACAG | 180 |
| GTAAAATGCT | GCACAGAGCA | TTCAGCGTTT | TTCTATTCAA | TTCAAAATAC | GAGTTACTTC | 240 |
| TTCAGCAACG | GTCTGCAACC | AAGGTGACAT | TTCCTTTAGT | ATGGACCAAC | ACCTGTTGCA | 300 |
| GCCATCCACT | CTACAGAGAA | TCCGAGCTTG | TTCCCGAAAA | CGCCCTTGGA | GTAAGAAATG | 360 |
| CTGCACAGAG | GAAGCTGTTG | GATGAACTCG | GTATCCCTGC | TGAAGATGTT | CCCGTTGATC | 420 |
| AGTTTACTCC | TTTAGGTCGC | ATGCTCTACA | AGGCTCCATC | TGATGGAAAG | TGGGGAGAAC | 480 |
| ATGAACTTGA | CTACCTACTT | TTCATAGTGA | GAGACGTTGC | TGTAAACCCG | AACCCAGATG | 540 |
| AAGTGGCGGA | TATCAAATAT | GTGACCAGAA | GAGTTAAAGG | AGCTGCTAAG | GAAAGCAGAT | 600 |
| GCGGGGGAGG | AGGGTTTGAA | GCTGTCTCCA | TGGTTCAGGT | TAGTGGTTGA | TAACTTCTTG | 660 |
| TTCAAGTGGT | GGGATCATGT | GCAAAAGGGT | ACACTCACTG | AAGCAATTGA | TATGAAAACC | 720 |
| ATACACAAGC | TGATATAGAA | ACACACCCTC | AACCGAAAAG | TTCAAGCCTA | ATAATTCGGG | 780 |
| TTGGGTCGGG | TCTACCATCA | ATTGTTTTTT | TCTTTTAAGA | AGTTTTAATC | TCTATTTGAG | 840 |
| CATGTTGATT | CTTGTCTTTT | GTGTGTAAGA | TTTTGGGTTT | CGTTTCAGTT | GTAATAATGA | 900 |
| ACCATTGATG | GTTTGCAATT | TCAAGTTCCT | ATCGACATGT | AGTGATCTAA | AAAA | 954 |
| <210> 35 <211> 1031 <212> DNA <213> Oryza | a sativa | | | | | |
| <400> 35 CCTCCCTTTG | CCTCGCGCAG | AGGCGGCCGC | GCCTTCTCCG | CCGCGAGGAT | GCCGGCGCC | 60 |
| GCCGCCGCCG | TGGAGGACGC | CGGGATGGAC | GAGGTCCAGA | AGCGGCTCAT | GTTCGACGAC | 120 |
| GAATGCATTT | TGGTGGATGA | ACAAGACAAT | GTTGTTGGCC | ATGAATCAAA | ATATAACTGC | 180 |
| CATCTGATGG | AAAAAATCGA | ATCTGAAAAT | CTACTTCATA | GGGCTTTCAG | TGTATTCCTG | 240 |



| _ | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| TTCAACTCAA | AATATGAACT | CCTACTCCAG | CAACGATCTG | CAACAAAGGT | TACATTTCCT | 300 |
| CTAGTTTGGA | CCAACACTTG | CTGCAGCCAT | CCTCTGTACC | GTGAGTCTGA | GCTTATACAG | 360 |
| GAAAACTACC | TTGGTGTTAG | AAATGCTGCT | CAGAGGAAGC | TCTTGGATGA | GCTGGGCATC | 420 |
| CCAGCTGAAG | ATGTGCCAGT | TGACCAATTC | ACCCCTCTTG | GTCGGATGCT | TTACAAGGCC | 480 |
| CCATCTGATG | GAAAATGGGG | TGAACACGAG | CTTGACTACC | TGCTGTTCAT | CGTCCGCGAC | 540 |
| GTGAAGGTAG | TCCCGAACCC | GGACGAAGTG | GCCGATGTGA | AATACGTGAG | CCGTGAGCAG | 600 |
| CTGAAGGAGC | TCATCCGCAA | AGCGGACGCC | GGAGAGGAAG | GCCTGAAGCT | GTCTCCCTGG | 660 |
| TTCCGGCTGG | TTGTTGACAA | CTTCCTCATG | GGCTGGTGGG | ATCACGTCGA | GAAAGGCACC | 720 |
| CTCAACGAGG | CCGTGGACAT | GGAGACCATC | CACAAGCTGA | AGTAAGGACT | GCGATGTTGT | 780 |
| GGCTGGAAAG | AATGATCCTG | AAGACTCTGT | TCTTGTGCTG | CTGCATATTA | CTCTTACCAG | 840 |
| GGAAGTTGCA | GAAGTCAGAA | GAAGCTTTTG | TATGTTTCTG | GGTTTGGAGC | TTGGAAGTGT | 900 |
| TGGGCTCTGC | TGACTGAGAG | ATTCCCTTAT | AGAGTGTCTA | TGTTAATTTA | GCAAACTTCT | 960 |
| ATATTATACA | TGATTAGTTA | ATTGTTCGGT | GTCTGAATAA | AGAACAATAG | CATGTTCCAT | 1020 |
| GTTTATTTGC | T | | | | | 1031 |
| <210> 36 | | | | | | |

<211> 232

<212> PRT

<213> Tagetes erecta

<400> 36

Met Gly Asp Asp Ser Gly Met Asp Ala Val Gln Arg Arg Leu Met Phe
1 10 15

Asp Asp Glu Cys Ile Leu Val Asp Glu Cys Asp Asn Val Val Gly His

Asp Thr Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Thr Gly Lys
35 40 45

Met Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu 50 55 60

Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro Leu Val
65 70 75 80

Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser Glu Leu 85 90 95

Val Pro Glu Asn Ala Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu 100 105 110

Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp Gln Phe 115 120 125

Thr Pro Leu Gly Arg Met Leu Tyr Lys Ala Pro Ser Asp Gly Lys Trp 130 135 140



| Gly Glu 145 | His | Glu | Asp 150 | Leu | Leu | Phe | Ile 155 | Val | Arg | Asp | Val | Ala 160 |
|----------------|-----|-----|------------|-----|-----|-----|------------|-----|-----|-----|-----|------------|
| | | | | | | | | | | | | |

Val Asn Pro Asn Pro Asp Glu Val Ala Asp Ile Lys Tyr Val Ser His 165 170 175

Glu Glu Leu Lys Glu Leu Leu Arg Lys Ala Asp Ala Gly Glu Glu Gly 180 185 190

Leu Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe Leu Phe 195 200 205

Lys Trp Trp Asp His Val Gln Lys Gly Thr Leu Thr Glu Ala Ile Asp 210 215 220

Met Lys Thr Ile His Lys Leu Ile 225 230

<210> 37

<211> 280

<212> PRT

<213> Lactuca Sativa

<400> 37

Met Leu Lys Phe Pro Pro Phe Lys Thr Ile Ala Thr Met Ile Ser Ser 1 5 10 15

Pro Tyr Ser Ser Phe Leu Leu Pro Arg Lys Ser Ser Phe Pro Pro Met 20 25 30

Pro Ser Leu Ala Ala Ser Val Phe Leu His Pro Leu Ser Ser Ala 35 40 45

Ala Met Gly Asp Ser Ser Met Asp Ala Val Gln Arg Arg Leu Met Phe 50 55 60

Asp Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Lys Val Val Gly His 65 70 75 80

Asp Thr Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Lys Gly Asn 85 90 95

Met Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu 100 105 110

Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro Leu Val 115 120 125

Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser Glu Leu 130 135 140

Ile Asp Glu Asn Ala Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu 145 150 155 160

Leu Asp Glu Leu Gly Ile Pro Gly Ala Asp Val Pro Val Asp Glu Phe 165 170 175

Thr Pro Leu Gly Arg Ile Leu Tyr Lys Ala Ala Ser Asp Gly Lys Trp
180 185 190

Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Met Val Arg Asp Val Gly

<u>.</u>.

195 200

Leu Asp Pro Asn Pro Asp Glu Val Lys Asp Val Lys Tyr Val Asn Arg 210 215 220

Glu Glu Leu Lys Glu Leu Val Arg Lys Ala Asp Ala Gly Glu Glu Gly 225 230 235 240

Val Lys Leu Ser Pro Trp Phe Lys Leu Ile Val Asp Asn Phe Leu Phe 245 250 255

Gln Trp Trp Asp Arg Leu His Lys Gly Thr Leu Thr Glu Ala Ile Asp 260 265 270

Met Lys Thr Ile His Lys Leu Thr 275 280

<210> 38

<211> 229

<212> PRT

<213> Lactuca Sativa

<400> 38

Met Gly Asp Asp Ser Gly Met Asp Ala Val Gln Arg Arg Leu Met Phe 1 5 10 15

Asp Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Asn Val Leu Gly His 20 25 30

Asp Thr Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Lys Asp Asn 35 40 45

Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu 50 55 60

Leu Leu Cln Gln Arg Ser Glu Thr Lys Val Thr Phe Pro Leu Val
65 70 75 80

Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser Glu Leu 85 90 95

Ile Pro Glu Asn Ala Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu 100 105 110

Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp Glu Phe 115 120 125

Thr Thr Leu Gly Arg Met Leu Tyr Lys Ala Pro Ser Asp Gly Lys Trp 130 135 140

Gly Glu His Glu Val Asp Tyr Leu Leu Phe Leu Val Arg Asp Val Ala 145 150 155 160

Val Asn Pro Asn Pro Asp Glu Val Ala Asp Ile Arg Tyr Val Asn Gln
165 170 175

Glu Glu Leu Lys Glu Leu Leu Arg Lys Ala Asp Ala Gly Glu Gly 180

Leu Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe Leu Phe 195 200 205



Lys Trp Trp Asp His Val Gln Lys Gly Thr Leu Asn Glu Ala Ile Asp 210 220

Met Lys Thr Ile His 225

<210> 39

<211> 295

<212> PRT

<213> Adonis Palaestina

<400> 39

Met Ser Ser Ile Arg Ile Asn Pro Leu Tyr Ser Ile Phe Ser Thr Thr 1 5 10 15

Thr Lys Thr Leu Ser Ala Ser Cys Ser Ser Pro Ala Val His Leu Gln
20 25 30

Gln Arg Cys Arg Thr Leu Ser Ile Ser Ser Ser Ile Thr Asn Ser Pro 35 40 45

Arg Arg Gly Leu Asn Arg Leu Phe Ala Ser Thr Ser Thr Met Gly Glu
50 60

Val Ala Asp Ala Gly Met Asp Ala Val Gln Lys Arg Leu Met Phe Asp 65 70 75 80

Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Lys Val Val Gly Tyr Asp 85 90 95

Ser Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala Glu Asn Leu 100 105 110

Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu Leu 115 120 125

Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro Leu Val Trp 130 135 140

Thr Asn Thr Cys Cys Ser His Pro Leu Phe Arg Asp Ser Glu Leu Ile 145 150 155 160

Glu Glu Asn Phe Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu Leu 165 170 175

Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp Glu Phe Thr 180 185 190

Pro Leu Gly Arg Ile Leu Tyr Lys Ala Pro Ser Asp Gly Lys Trp Gly 195 200 205

Glu His Glu Leu Asp Tyr Leu Leu Phe Ile Val Arg Asp Val Lys Tyr 210 215 220

Asp Pro Asn Pro Asp Glu Val Ala Asp Ala Lys Tyr Val Asn Arg Glu 225 230 235 240

Glu Leu Lys Glu Ile Leu Arg Lys Ala Asp Ala Gly Glu Glu Gly Ile 245 250 . 255

Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe Leu Phe Lys

265

270

Trp Trp Asp His Val Glu Glu Gly Lys Ile Lys Asp Val Ala Asp Met 275 280 285

Lys Thr Ile His Lys Leu Thr 290 295

<210> 40

<211> 234

<212> PRT

<213> Adonis Palaestina

<400> 40

Met Gly Glu Val Thr Asp Ala Gly Met Asp Ala Val Gln Lys Arg Leu 1 5 10 15

Met Phe Asp Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Lys Val Val 20 25 30

Gly His Asp Ser Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala $35 \hspace{1cm} 40 \hspace{1cm} 45$

Glu Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys
50 55 60

Tyr Glu Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro 65 70 75 80

Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu Phe Arg Asp Ser 85 90 95

Glu Leu Ile Glu Glu Asn Tyr Leú Gly Val Arg Asn Ala Ala Gln Arg 100 105 110

Lys Leu Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp 115 120 125

Glu Phe Thr Pro Leu Gly Arg Ile Leu Tyr Lys Ala Pro Ser Asp Gly 130 135 140

Lys Trp Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Ile Val Arg Asp 145 150 155 160

Val Lys Tyr Asp Pro Asp Pro Asp Glu Val Ala Asp Ala Lys Tyr Val 165 170 175

Asn Arg Glu Glu Leu Arg Glu Ile Leu Arg Lys Ala Asp Ala Gly Glu 180 185 190

Glu Gly Leu Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe 195 200 205

Leu Phe Lys Trp Trp Asp His Val Glu Gln Gly Thr Ile Lys Glu Val 210 215 220

Ala Asp Met Lys Thr Ile His Lys Leu Thr 225 230

<210> 41 <211> 238



<212> PRT <213> Oryza Sativa

<400> 41 Met Ala Gly Ala Ala Ala Val Glu Asp Ala Gly Met Asp Glu Val

Gln Lys Arg Leu Met Phe Asp Asp Glu Cys Ile Leu Val Asp Glu Gln

Asp Asn Val Val Gly His Glu Ser Lys Tyr Asn Cys His Leu Met Glu

Lys Ile Glu Ser Glu Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu

Phe Asn Ser Lys Tyr Glu Leu Leu Gln Gln Arg Ser Ala Thr Lys 75

Val Thr Phe Pro Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu

Tyr Arg Glu Ser Glu Leu Ile Gln Glu Asn Tyr Leu Gly Val Arg Asn 105

Ala Ala Gln Arg Lys Leu Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp 120

Val Pro Val Asp Gln Phe Thr Pro Leu Gly Arg Met Leu Tyr Lys Ala 130

Pro Ser Asp Gly Lys Trp Gly Glu His Glu Leu Asp Tyr Leu Leu Phe 150

Ile Val Arg Asp Val Lys Val Val Pro Asn Pro Asp Glu Val Ala Asp

Val Lys Tyr Val Ser Arg Glu Gln Leu Lys Glu Leu Ile Arg Lys Ala 180

Asp Ala Gly Glu Glu Gly Leu Lys Leu Ser Pro Trp Phe Arg Leu Val 200

Val Asp Asn Phe Leu Met Gly Trp Trp Asp His Val Glu Lys Gly Thr 210

Leu Asn Glu Ala Val Asp Met Glu Thr Ile His Lys Leu Lys

<210> 42

<211> 233

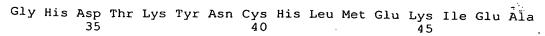
<212> PRT

<213> Arabidopsis thaliana

<400> 42

Met Thr Asp Ser Asn Asp Ala Gly Met Asp Ala Val Gln Arg Arg Leu

Met Phe Glu Asp Glu Cys Ile Leu Val Asp Glu Asn Asn Arg Val Val 20



Glu Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys
50 55 60

Tyr Glu Leu Leu Gln Gln Arg Ser Lys Thr Lys Val Thr Phe Pro 65 70 75 80

Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser 85 90 95

Glu Leu Ile Glu Glu Asn Val Leu Gly Val Arg Asn Ala Ala Gln Arg 100 105 110

Lys Leu Phe Asp Glu Leu Gly Ile Val Ala Glu Asp Val Pro Val Asp 115 120 125

Glu Phe Thr Pro Leu Gly Arg Met Leu Tyr Lys Ala Pro Ser Asp Gly 130 135 140

Lys Trp Gly Glu His Glu Val Asp Tyr Leu Leu Phe Ile Val Arg Asp 145 150 155 160

Val Lys Leu Gln Pro Asn Pro Asp Glu Val Ala Glu Ile Lys Tyr Val 165 170 175

Ser Arg Glu Glu Leu Lys Glu Leu Val Lys Lys Ala Asp Ala Gly Asp 180 185 190

Glu Ala Val Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe 195 200 205

Leu Met Lys Trp Trp Asp His Val Glu Lys Gly Thr Ile Thr Glu Ala 210 215 220

Ala Asp Met Lys Thr Ile His Lys Leu 225 230

<210> 43

<211> 293

<212> PRT

<213> Haematococcus pluvialis

<400> 43

Met Leu Arg Ser Leu Leu Arg Gly Leu Thr His Ile Pro Arg Val Asn
1 5 10 15

Ser Ala Gln Gln Pro Ser Cys Ala His Ala Arg Leu Gln Phe Lys Leu 20 25 30

Arg Ser Met Gln Leu Leu Ser Glu Asp Arg Thr Asp His Met Arg Gly 35 40 45

Ala Ser Thr Trp Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys
50 55 60

Asp Glu Cys Ile Leu Val Asp Val Glu Asp Asn Ile Thr Gly His Ala 65 70 75 80

Ser Lys Leu Glu Cys His Lys Phe Leu Pro His Gln Pro Ala Gly Leu





Leu His Arg Ala Phe Ser Val Phe Leu Phe Asp Asp Gln Gly Arg Leu 100 105 110

Leu Leu Gln Gln Arg Ala Arg Ser Lys Ile Thr Phe Pro Ser Val Trp 115 120 125

Thr Asn Thr Cys Cys Ser His Pro Leu His Gly Gln Thr Pro Asp Glu 130 135 140

Val Asp Gln Leu Ser Gln Val Ala Asp Gly Thr Val Pro Gly Ala Lys
145 150 155 160

Ala Ala Ile Arg Lys Leu Glu His Glu Leu Gly Ile Pro Ala His 165 170 175

Gln Leu Pro Ala Ser Ala Phe Arg Phe Leu Thr Arg Leu His Tyr Cys 180 185 190

Ala Ala Asp Val Gln Pro Ala Ala Thr Gln Ser Ala Leu Trp Gly Glu 195 200 205

His Glu Met Asp Tyr Ile Leu Phe Ile Arg Ala Asn Val Thr Leu Ala 210 225 220

Pro Asn Pro Asp Glu Val Asp Glu Val Arg Tyr Val Thr Gln Glu Glu 225 235 240

Leu Arg Gln Met Met Gln Pro Asp Asn Gly Leu Gln Trp Ser Pro Trp 245 250 255

Phe Arg Ile Ile Ala Ala Arg Phe Leu Glu Arg Trp Trp Ala Asp Leu 260 265 270

Asp Ala Ala Leu Asn Thr Asp Lys His Glu Asp Trp Gly Thr Val His 275 280 285

His Ile Asn Glu Ala 290

<210> 44

<211> 304

<212> PRT

<213> Haematococcus pluvialis

<400> 44

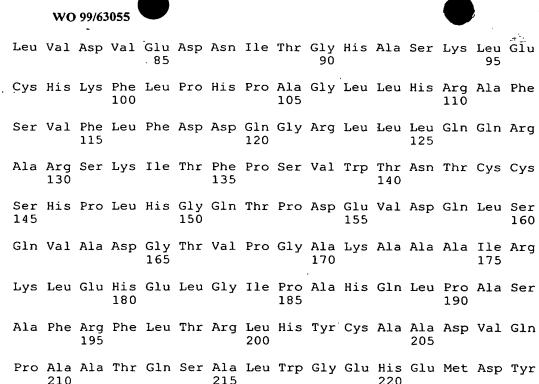
Met Leu Arg Ser Leu Leu Arg Gly Leu Thr His Ile Pro Arg Val Asn
1 5 10 15

Ser Ala Gln Gln Pro Ser Cys Ala His Ala Arg Leu Gln Phe Lys Leu 20 25 30

Arg Ser Met Gln Met Thr Leu Met Gln Pro Ser Ile Ser Ala Asn Leu 35 40 45

Ser Arg Ala Glu Asp Arg Thr Asp His Met Arg Gly Ala Ser Thr Trp 50 55 60

Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys Asp Glu Cys Ile 65 70 75 80



Ile Leu Phe Ile Arg Ala Asn Val Thr Leu Ala Pro Asn Pro Asp Glu 225 230 235 240

Val Asp Glu Val Arg Tyr Val Thr Gln Glu Glu Leu Arg Gln Met Met 245 250 255

Gln Pro Asp Asn Gly Leu Gln Trp Ser Pro Trp Phe Arg Ile Ile Ala 260 265 270

Ala Arg Phe Leu Glu Arg Trp Trp Ala Asp Leu Asp Ala Ala Leu Asn 275 280 285

Thr Asp Lys His Glu Asp Trp Gly Thr Val His His Ile Asn Glu Ala 290 295 300

<210> 45

<211> 307

<212> PRT

<213> Chlamydomonas reinhardtii

<400> 45

Met Arg Ser Ser Phe Ile Glu Pro Lys Pro Arg Ala Gln Pro Val Leu 1 5 10 15

Ser Arg Gly Arg Ala Ser Met Arg Leu Ala Gln Ser Arg Ala Leu Val 20 25 30

Ala Arg Val Ser Ser Ala Leu Trp Pro Gly Ala Gly Leu Ser Gln Ala 35 40 45

Gln Ser Val Ala Val Arg Met Ala Ser Ser Ser Thr Trp Glu Gly Thr

Gly Leu Ser Gln Asp Asp Phe Met Gln Arg Asp Glu Cys Leu Val Val

65



70

| | · Control of the cont | |
|---|--|--|
| | | |
| i | | |
| 1 | | |
| | | |
| | | |

.75

Asp Glu Gln Asp Arg Leu Leu Gly Thr Ala Asn Lys Tyr Asp Cys His 95

Arg Phe Glu Ala Ala Lys Gly Gln Pro Cys Gly Arg Leu His Arg Ala 100

Phe Ser Val Phe Leu Phe Ser Pro Asp Gly Arg Leu Leu Cys Gly Gla

Phe Ser Val Phe Leu Phe Ser Pro Asp Gly Arg Leu Leu Gln Gln 115 120 125

Arg Ala Ala Ser Lys Val Thr Phe Pro Gly Val Trp Thr Asn Thr Cys 130 140

Cys Ser His Pro Leu Ala Gly Gln Ala Pro Asp Glu Val Asp Leu Pro 145 150 155 160

Ala Ala Val Ala Ser Gly Gln Val Pro Gly Ile Lys Ala Ala Ala Val 165 170 175

Arg Lys Leu Gln His Glu Leu Gly Ile Pro Pro Glu Gln Val Pro Ala 180 185 190

Ser Ser Phe Ser Phe Leu Thr Arg Leu His Tyr Cys Ala Ala Asp Thr 195 200 205

Ala Thr His Gly Pro Ala Ala Glu Trp Gly Glu His Glu Val Asp Tyr 210 215 220

Val Leu Phe Val Arg Pro Gln Gln Pro Val Ser Leu Gln Pro Asn Pro 225 230 235 240

Asp Glu Val Asp Ala Thr Arg Tyr Val Thr Leu Pro Glu Leu Gln Ser 245 250 255

Met Met Ala Asp Pro Gly Leu Ser Trp Ser Pro Trp Phe Arg Ile Leu 260 265 270

Ala Thr Gln Pro Ala Phe Leu Pro Ala Trp Trp Gly Asp Leu Lys Arg 275 280 285

Arg Trp Arg Pro Gly Gly Ser Arg Leu Ser Asp Trp Gly Thr Ile His 290 295 300

Arg Val Met 305

<210> 46

<211> 1848

<212> DNA

<213> Adonis palaestina

<400> 46

WO 99/63055

| *************************************** | 05055 | | | | | |
|---|------------|------------|------------|------------|------------|------|
| TTTTCAAGTG | AGGGCTGATG | GTGGAAGCGG | GAGTAGAACT | TCTGTTGCTT | ATAAAGĀGGG | 300 |
| TTTTGTGGAC | GAGGAGGATT | TTATCAAAGC | TGGTGGTTCT | GAGCTTTTGT | TTGTCCAAAT | 360 |
| GCAGCAAACA | AAGTCTATGG | AGAAACAGGC | CAAGCTCGCC | GATAAGTTGC | CACCAATACC | 420 |
| TTTCGGAGAA | TCTGTGATGG | ACTTGGTTGT | AATAGGTTGT | GGACCTGCTG | GTCTTTCACT | 480 |
| GGCTGCAGAA | GCTGCTAAGC | TAGGCTTGAA | AGTTGGCCTT | ATTGGTCCTG | ATCTTCCTTT | 540 |
| TACAAATAAT | TATGGTGTGT | GGGAAGACGA | GTTCAAAGAT | CTTGGACTTG | AACGTTGTAT | 600 |
| CGAGCATGCT | TGGAAGGACA | CCATCGTATA | TCTTGACAAT | GATGCTCCTG | TCCTTATTGG | 660 |
| TCGTGCATAT | GGACGAGTTA | GCCGGCATTT | GCTGCATGAA | GAGTTGCTGA | AAAGGTGTGT | 720 |
| CGAGTCAGGT | GTATCATATC | TGAATTCTAA | AGTGGAAAGG | ATCACTGAAG | CTGGTGATGG | 780 |
| CCATAGTCTT | GTAGTTTGTG | AAAACGACAT | CTTTATCCCT | TGCAGGCTTG | CTACTGTTGC | 840 |
| ATCTGGAGCA | GCTTCAGGGA | AACTTTTGGA | GTATGAAGTA | GGTGGCCCTC | GTGTTTGTGT | 900 |
| CCAAACTGCT | TATGGTGTGG | AGGTTGAGGT | GGAGAACAAT | CCATACGATC | CCAACTTAAT | 960 |
| GGTATTTATG | GACTACAGAG | ACTATATGCA | ACAGAAATTA | CAGTGCTCGG | AAGAAGAATA | 1020 |
| TCCAACATTT | CTCTATGTCA | TGCCCATGTC | GCCAACAAGA | CTTTTTTTTG | AGGAAACCTG | 1080 |
| TTTGGCCTCA | AAAGATGCCA | TGCCTTTCGA | TCTACTGAAG | AGAAAACTAA | TGTCACGATT | 1140 |
| GAAGACTCTG | GGTATCCAAG | TTACAAAAAT | TTATGAAGAG | GAATGGTCTT | ATATTCCTGT | 1200 |
| TGGGGGTTCT | TTACCAAACA | CAGAGCAAAA | GAACCTAGCA | TTTGGTGCTG | CAGCAAGCAT | 1260 |
| GGTGCATCCA | GCAACAGGCT | ATTCGGTTGT | ACGATCACTA | TCAGAAGCTC | CAAAATATGC | 1320 |
| TTCTGTAATT | GCAAAGATTT | TGAAGCAAGA | TAACTCTGCA | TATGTGGTTT | CTGGACAAAG | 1380 |
| CAGTGCAGTA | AACATTTCAA | TGCAAGCATG | GAGCAGTCTT | TGGCCAAAGG | AGCGAAAACG | 1440 |
| TCAAAGAGCA | TTCTTTCTTT | TCGGGTTAGA | GCTTATTGTG | CAGCTAGATA | TTGAAGCAAC | 1500 |
| CAGAACGTTC | TTTAGAACCT | TCTTCCGCTT | GCCAACTTGG | ATGTGGTGGG | GTTTCCTTGG | 1560 |
| GTCTTCACTA | TCATCTTTCG | ATCTTGTATT | GTTTTCCATG | TACATGTTTG | TTTTGGCCCC | 1620 |
| GAACAGCATG | AGGATGTCAC | TTGTGAGACA | TTTGCTTTCA | GATCCTTCTG | GTGCAGTTAT | 1680 |
| GGTTAAAGCT | TACCTCGAAA | GGTAATCTGT | TTTATGAAAC | TATAGTGTCT | САТТАААТАА | 1740 |
| ATGAGGATCC | TTCGTATATG | TATATGATCA | TCTCTATGTA | TATCCTATAT | TCTAATCTCA | 1800 |
| TAAAGTAATC | GAAAATTCAT | TGATAGAAAA | АААААААА | AAAAAAA | | 1848 |

PCT/US99/12121

Met Glu Leu Gly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp

<210> 47

<211> 529

<212> PRT

<213> Adonis palaestina

<400> 47

10 15 Thr Phe Gly Thr Arg Asn Leu Ser Ser Ser Lys Leu Ala Tyr Asn Ile His Arg Tyr Gly Ser Ser Cys Arg Val Asp Phe Gln Val Arg Ala Asp Gly Gly Ser Gly Ser Arg Ser Ser Val Ala Tyr Lys Glu Gly Phe Val Asp Glu Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu Leu Leu Phe Val Gln Met Gln Gln Thr Lys Ser Met Glu Lys Gln Ala Lys Leu Ala Asp Lys Leu Pro Pro Ile Pro Phe Gly Glu Ser Val Met Asp Leu Val Val 105 Ile Gly Cys Gly Pro Ala Gly Leu Ser Leu Ala Ala Glu Ala Ala Lys Leu Gly Leu Lys Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Glu Arg 155 Cys Ile Glu His Ala Trp Lys Asp Thr Ile Val Tyr Leu Asp Asn Asp Ala Pro Val Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu His Glu Glu Leu Leu Lys Arg Cys Val Glu Ser Gly Val Ser Tyr 200 Leu Asp Ser Lys Val Glu Arg Ile Thr Glu Ala Gly Asp Gly His Ser Leu Val Val Cys Glu Asn Glu Ile Phe Ile Pro Cys Arg Leu Ala Thr 225 Val Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Glu Tyr Glu Val Gly 250 Gly Pro Arg Val Cys Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Glu Asn Asn Pro Tyr Asp Pro Asn Leu Met Val Phe Met Asp Tyr Arg 280 Asp Tyr Met Gln Gin Lys Leu Gln Cys Ser Glu Glu Glu Tyr Pro Thr 295 Phe Leu Tyr Val Met Pro Met Ser Pro Thr Arg Leu Phe Phe Glu Glu 315 Thr Cys Leu Ala Ser Lys Asp Ala Met Pro Phe Asp Leu Leu Lys Arg

330



Lys Leu Met Ser Arg Leu Lys Thr Leu Gly Ile Gln Val Thr Lys Val 345 Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys 390 Tyr Ala Ser Val Ile Ala Lys Ile Leu Lys Gln Asp Asn Ser Ala Tyr Val Val Ser Gly Gln Ser Ser Ala Val Asn Ile Ser Met Gln Ala Trp Ser Ser Leu Trp Pro Lys Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu Phe Gly Leu Glu Leu Ile Val Gln Leu Asp Ile Glu Ala Thr Arg Thr Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp Trp Gly Phe Leu Gly Ser Ser Leu Ser Ser Phe Asp Leu Val Leu Phe Ser Met Tyr

Met Phe Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His 500 505

Leu Leu Ser Asp Pro Ser Gly Ala Val Met Val Arg Ala Tyr Leu Glu

Arg

<210> 48

<211> 378

<212> PRT

<213> Potato

<400> 48

Asp Glu Phe Lys Asp Leu Gly Leu Gln Ala Cys Ile Glu His Val Trp

Arg Asp Thr Ile Val Tyr Leu Asp Asp Asp Pro Ile Leu Ile Gly

Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu His Glu Glu Leu Leu

Lys Arg Cys Val Glu Ala Gly Val Leu Tyr Leu Asn Ser Lys Val Asp

Arg Ile Val Glu Ala Thr Asn Gly His Ser Leu Val Glu Cys Glu Gly

Asp Val Val Ile Pro Cys Arg Phe Val Thr Val Ala Ser Gly Ala Ala



| Ser G | ly I | ys | Phe | Leu | Gln | Tyr | Glu | Leu | Gly | Gly | Pro | Arg | Val | Ser | Val |
|-------|------|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | | | 100 | | | | | 105 | | | | | 110 | | |

Gln Thr Ala Tyr Gly Val Glu Val Glu Val Asp Asn Asn Pro Phe Asp 115 120 125

Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp Tyr Val Arg His Asp 130 135 140

Ala Gln Ser Leu Glu Ala Lys Tyr Pro Thr Phe Leu Tyr Ala Met Pro 145 150 155 160

Met Ser Pro Thr Arg Val Phe Phe Glu Glu Thr Cys Leu Ala Ser Lys 165 170 175

Asp Ala Met Pro Phe Asp Leu Leu Lys Lys Leu Met Leu Arg Leu 180 185 190

Asn Thr Leu Gly Val Arg Ile Lys Glu Ile Tyr Glu Glu Glu Trp Ser 195 200 205

Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Thr Leu 210 215 220

Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser 225 230 235 240

Val Val Arg Ser Leu Ser Glu Ala Pro Lys Cys Ala Phe Val Leu Ala 245 250 255

Asn Ile Leu Arg Gln Asn His Ser Lys Asn Met Leu Thr Ser Ser Ser 260 265 270

Thr Pro Ser Ile Ser Thr Gln Ala Trp Asn Thr Leu Trp Pro Gln Glu 275 280 285

Arg Lys Arg Gln Arg Ser Phe Phe Leu Phe Gly Leu Ala Leu Ile Leu 290 295 300

Gln Leu Asp Ile Glu Gly Ile Arg Ser Phe Phe Arg Ala Phe Phe Arg 305 . 310 315 320

Val Pro Lys Trp Met Trp Gln Gly Phe Leu Gly Ser Ser Leu Ser Xaa 325 330 335

Ala Asp Leu Met Leu Phe Ala Phe Tyr Met Phe Ile Ile Ala Pro Asn 340 345 350

Asp Met Arg Arg Gly Leu Ile Arg His Leu Leu Ser Asp Pro Thr Gly 355 360 365

Ala Thr Leu Ile Arg Thr Tyr Leu Thr Phe 370

<210> 49

<211> 524

<212> PRT

<213> Arabidopsis thaliana

<400> 49



| Met 1 | Glu | Cys | Val | Gly 5 | Ala | Arg | Asn | Phe | Ala 10 | Ala | Met | Ala | Val | Ser 15 | Thr |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Phe | Pro | Ser | Trp 20 | Ser | Cys | Arg | Arg | Lys 25 | Phe | Pro | Val | Val | Lys 30 | Arg | Tyr |
| Ser | Tyr | Arg 35 | Asn | Ile | Arg | Phe | Gly 40 | Leu | Cys | Ser | Val | Arg 45 | Ala | Ser | Gly |
| Gly | Gly 50 | Ser | Ser | Gly | Ser | Glu 55 | Ser | Cys | Val | Ala | Val 60 | Arg | Glu | Asp | Phe |
| Ala 65 | Asp | Glu | Glu | Asp | Phe 70 | Val | Lys | Ala | Gly | Gly 75 | Ser | Glu | Ile | Leu | Phe 80 |
| Val | Gln | Met | Gln | Gln 85 | Asn | Lys | Asp | Met | Asp 90 | Glu | Gln | Ser | Lys | Leu 95 | Val |
| Asp | Lys | Leu | Pro 100 | Pro | Ile | Ser | Ile | Gly 105 | Asp | Gly | Ala | Leu | Asp 110 | His | Val |
| Val | Ile | Gly 115 | Cys | Gly | Pro | Ala | Gly 120 | Leu | Ala | Leu | Ala | Ala 125 | Glu | Ser | Ala |
| Lys | Leu 130 | Gly | Leu | Lys | Val | Gly 135 | Leu | Ile | Gly | Pro | Asp 140 | Leu | Pro | Phe | Thr |
| Asn 145 | Asn | Tyr | Gly | Val | Trp 150 | Glu | Asp | Glu | Phe | Asn 155 | Asp | Leu | Gly | Leu | Gln 160 |
| Lys | Cys | Ile | Glu | His 165 | Val | Trp | Arg | Glu | Thr 170 | Ile | Val | Tyr | Leu | Asp 175 | |
| Asp | Lys | Pro | Ile 180 | Thr | Ile | Gly | Arg | Ala 185 | Tyr | Gly | Arg | Val | Ser 190 | Arg | Arg |
| Leu | Leu | His 195 | Glu | Glu | Leu | Leu | Arg 200 | Arg | Cys | Val | Glu | Ser 205 | Gly | Val | Ser |
| Tyr | Leu 210 | Ser | Ser | Lys | Val | Asp 215 | Ser | Ile | Thr | Glu | Ala 220 | Ser | Asp | Gly | Leu |
| Arg 225 | Leu | Val | Ala | Cys | Asp 230 | Asp | Asn | Asn | Val | Ile 235 | Pro | Cys | Arg | Leu | Ala 240 |
| Thr | Val | Ala | Ser | Gly 245 | Ala | Ala | Ser | Gly | Lys 250 | Leu | Leu | Gln | Tyr | Glu 255 | Val |
| Gly | Gly | Pro | Arg 260 | Val | Cys | Val | Gln | Thr 265 | Ala | Tyr | Gly | Val | Glu 270 | Val | Glu |
| Val | Glu | Asn 275 | Ser | Pro | Tyr | Asp | Pro 280 | qzA | Gln | Met | Val | Phe 285 | Met | Asp | Tyr |
| Arg | Asp 290 | Tyr | Thr | Asn | Glu | Lys 295 | Val | Arg | Ser | Leu | Glu 300 | Ala | Glu | Tyr | Pro |
| Thr 305 | Phe | Leu | Tyr | Ala | Met 310 | Pro | Met | Thr | Lys | Ser 315 | Arg | Leu | Phe | Phe | Glu 320 |
| Glu | Thr | Cys | Leu | Ala | Ser | Lys | Asp | Val | Met | Pro | Phe | Asp | Leu | Leu | Lys |

325 330 335

Thr Lys Leu Met Leu Arg Leu Asp Thr Leu Gly Ile Arg Ile Leu Lys 340 345 350

Thr Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro 355 360 365

Asn Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val 370 380

His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro 385 390 395 400

Lys Tyr Ala Ser Val Ile Ala Glu Ile Leu Arg Glu Glu Thr Thr Lys
405 410 415

Gln Ile Asn Ser Asn Ile Ser Arg Gln Ala Trp Asp Thr Leu Trp Pro
420 425 430

Pro Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu Phe Gly Leu Ala Leu 435 440 445

Ile Val Gln Phe Asp Thr Glu Gly Ile Arg Ser Phe Phe Arg Thr Phe 450 455 460

Phe Arg Leu Pro Lys Trp Met Trp Gln Gly Phe Leu Gly Ser Thr Leu 465 470 475 480

Thr Ser Gly Asp Leu Val Leu Phe Ala Leu Tyr Met Phe Val Ile Ser 485 490 495

Pro Asn Asn Leu Arg Lys Gly Leu Ile Asn His Leu Ile Ser Asp Pro 500 505 510

Thr Gly Ala Thr Met Ile Lys Thr Tyr Leu Lys Val 515 520

<210> 50

<211> 529

<212> PRT

<213> Adonis palaestina

<400> 50

Met Glu Leu Leu Gly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp
1 5 10 15

Thr Phe Gly Thr Arg Asn Leu Ser Ser Ser Lys Leu Ala Tyr Asn Ile 20 25 30

His Arg Tyr Gly Ser Ser Cys Arg Val Asp Phe Gln Val Arg Ala Asp 35 40 45

Gly Gly Ser Gly Ser Arg Ser Ser Val Ala Tyr Lys Glu Gly Phe Val 50 60

Asp Glu Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu Leu Leu Phe Val 65 70 75 80

Gln Met Gln Gln Thr Lys Ser Met Glu Lys Gln Ala Lys Leu Ala Asp

95 Lys Leu Pro Pro Ile Pro Phe Gly Glu Ser Val Met Asp Leu Val Val 105 Ile Gly Cys Gly Pro Ala Gly Leu Ser Leu Ala Ala Glu Ala Ala Lys Leu Gly Leu Lys Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Glu Arg. Cys Ile Glu His Ala Trp Lys Asp Thr Ile Val Tyr Leu Asp Asn Asp Ala Pro Val Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu His Glu Glu Leu Leu Lys Arg Cys Val Glu Ser Gly Val Ser Tyr Leu Asp Ser Lys Val Glu Arg Ile Thr Glu Ala Gly Asp Gly His Ser Leu Val Val Cys Glu Asn Glu Ile Phe Ile Pro Cys Arg Leu Ala Thr 235 Val Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Glu Tyr Glu Val Gly Gly Pro Arg Val Cys Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val 260 Glu Asn Asn Pro Tyr Asp Pro Asn Leu Met Val Phe Met Asp Tyr Arg 280 Asp Tyr Met Gln Gln Lys Leu Gln Cys Ser Glu Glu Glu Tyr Pro Thr Phe Leu Tyr Val Met Pro Met Ser Pro Thr Arg Leu Phe Phe Glu Glu 315 Thr Cys Leu Ala Ser Lys Asp Ala Met Pro Phe Asp Leu Leu Lys Arg 330 Lys Leu Met Ser Arg Leu Lys Thr Leu Gly Ile Gln Val Thr Lys Val Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val His 375 Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys 395 Tyr Ala Ser Val Ile Ala Lys Ile Leu Lys Gln Asp Asn Ser Ala Tyr 410



Val Val Ser Gly Gln Ser Ser Ala Val Asn Ile Ser Met Gln Ala Trp
420 425 430

Ser Ser Leu Trp Pro Lys Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu 435 440 445

Phe Gly Leu Glu Leu Ile Val Gln Leu Asp Ile Glu Ala Thr Arg Thr 450 460

Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp Trp Gly Phe 465 470 475 480

Leu Gly Ser Ser Leu Ser Ser Phe Asp Leu Val Leu Phe Ser Met Tyr
485 490 495

Met Phe Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His 500 505 510

Leu Leu Ser Asp Pro Ser Gly Ala Val Met Val Arg Ala Tyr Leu Glu 515 520 525

Arg

<210> 51

<211> 529

<212> PRT

<213> Adonis palaestina

<400> 51

Met Glu Leu Leu Gly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp
1 5 10 15

Thr Phe Gly Thr Arg Asn Leu Ser Ser Ser Lys Leu Ala Tyr Asn Ile
20 25 30

His Arg Tyr Gly Ser Ser Cys Arg Val Asp Phe Gln Val Arg Ala Asp 35 40 45

Gly Gly Ser Gly Ser Arg Thr Ser Val Ala Tyr Lys Glu Gly Phe Val
50 60

Asp Glu Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu Leu Leu Phe Val 65 70 75 80

Gln Met Gln Gln Thr Lys Ser Met Glu Lys Gln Ala Lys Leu Ala Asp 85 90 95

Lys Leu Pro Pro Ile Pro Phe Gly Glu Ser Val Met Asp Leu Val Val 100 105 110

Ile Gly Cys Gly Pro Ala Gly Leu Ser Leu Ala Ala Glu Ala Ala Lys 115 120 125

Leu Gly Leu Lys Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn 130 135 140

Asn Tyr Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Glu Arg 145 150 155 160



| Cys | Ile | Glu | His | Ala 165 | Trp | Lys | Asp | Thr | Ile 170 | Val | Туr | Leu | Asp | Asn 175 | Asp |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Ala | Pro | Val | Leu 180 | Ile | Gly | Arg | Ala | Tyr 185 | Gly | Arg | Val | Ser | Arg 190 | His | Leu |
| Leu | His | Glu 195 | Glu | Leu | Leu | Lys | Arg 200 | Суѕ | Val | Glu | Ser | Gly 205 | Val | Ser | Tyr |
| Leu | Asn 210 | Ser | Lys | Val | Glu | Arg 215 | Ile | Thr | Glu | Ala | Gly 220 | Asp | Gly | His | Ser |
| Leu 225 | Val | Val | Cys | Glu | Asn 230 | Asp | Ile | Phe | Ile | Pro 235 | Cys | Arg | Leu | Ala | Thr 240 |
| Val | Ala | Ser | Gly | Ala 245 | Ala | Ser | Gly | Lys | Leu 250 | Leu | Glu | Tyr | Glu | Val 255 | Gly |
| Gly | Pro | Arg | Val 260 | Cys | Val | Gln | Thr | Ala 265 | Tyr | Gly | Val | Glu | Val 270 | Glu | Val |
| Glu | Asn | Asn 275 | Pro | Tyr | Asp | Pro | Asn 280 | Leu | Met | Val | Phe | Met 285 | Asp | Tyr | Arg |
| Asp | Tyr 290 | Met | Gln | Gln | Lys | Leu 295 | Gln | Cys | Ser | Glu | Glu 300 | Glu | Tyr | Pro | Thr |
| Phe 305 | Leu | Tyr | Val | Met | Pro 310 | Met | Ser | Pro | Thr | Arg 315 | Leu | Phe | Phe | Glu | Glu 320 |
| Thr | Cys | Leu | Ala | Ser 325 | Lys | Asp | Ala | Met | Pro 330 | Phe | Asp | Leu | Leu | Lys 335 | Arg |
| Lys | Leu | Met | Ser 340 | Arg | Leu | Lys | Thr | Leu 345 | Gly | Ile | Gln | Val | Thr 350 | Lys | Ile |
| Tyr | Glu | Glu 355 | Glu | Trp | Ser | Tyr | 11e 360 | Pro | Val | Gly | Gly | Ser 365 | Leu | Pro | Asn |
| Thr | Glu 370 | Gln | Lys | Asn | Leu | Ala 375 | Phe | Gly | Ala | Ala | Ala 380 | Ser | Met | Val | His |
| Pro 385 | Ala | Thr | Gly | Tyr | Ser 390 | Val | Val | Arg | Ser | Leu 395 | Ser | Glu | Ala | Pro | Lys 400 |
| Tyr | Ala | Ser | Val | Ile 405 | Ala | Lys | Ile | Leu | Lys 410 | Gln | Asp | Asn | Ser | Ala 415 | Tyr |
| Val | Val | Ser | Gly 420 | Gln | Ser | Ser | Ala | Val 425 | Asn | Ile | Ser | Met | Gln 430 | Ala | Trp |
| Ser | Ser | Leu 435 | Trp | Pro | Lys | Glu | Arg 440 | Lys | Arg | Gln | Arg | Ala 445 | Phe | Phe | Leu |
| Phe | Gly 450 | Leu | Glu | Leu | Ile | Val 455 | Gln | Leu | Asp | Ile | Glu 460 | Ala | Thr | Arg | Thr |
| Phe 465 | Phe | Arg | Thr | Phe | Phe 470 | Arg | Leu | Pro | Thr | Trp 475 | Met | Trp | Trp | .Gly | Phe 480 |
| Leu | Gly | Ser | Ser | Leu | Ser | Ser | Phe | Asp | Leu | Val | Leu | Phe | Ser | Met | Tyr |

490

495

Met Phe Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His 500 505 510

Leu Leu Ser Asp Pro Ser Gly Ala Val Met Val Lys Ala Tyr Leu Glu 515 520 525

Arg

<210> 52

<211> 533

<212> PRT

<213> Lettuce

<400> 52

Met Glu Cys Phe Gly Ala Arg Asn Met Thr Ala Thr Met Ala Val Phe
1 5 10 15

Thr Cys Pro Arg Phe Thr Asp Cys Asn Ile Arg His Lys Phe Ser Leu 20 25 30

Leu Lys Gln Arg Arg Phe Thr Asn Leu Ser Ala Ser Ser Ser Leu Arg
35 40 45

Gln Ile Lys Cys Ser Ala Lys Ser Asp Arg Cys Val Val Asp Lys Gln
50 60

Gly Ile Ser Val Ala Asp Glu Glu Asp Tyr Val Lys Ala Gly Gly Ser
65 70 75 80

Glu Leu Phe Phe Val Gln Met Gln Arg Thr Lys Ser Met Glu Ser Gln 85 90 95

Ser Lys Leu Ser Glu Lys Leu Ala Gln Ile Pro Ile Gly Asn Cys Ile 100 105 110

Leu Asp Leu Val Val Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala 115 120 125

Ala Glu Ser Ala Lys Leu Gly Leu Asn Val Gly Leu Ile Gly Pro Asp 130 135 140

Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Gln Asp Glu Phe Ile Gly
145 150 155 160

Leu Gly Leu Glu Gly Cys Ile Glu His Ser Trp Lys Asp Thr Leu Val 165 170 175

Tyr Leu Asp Asp Ala Asp Pro Ile Arg Ile Gly Arg Ala Tyr Gly Arg 180 185 190

Val His Arg Asp Leu Leu His Glu Glu Leu Leu Arg Arg Cys Val Glu 195 200 205

Ser Gly Val Ser Tyr Leu Ser Ser Lys Val Glu Arg Ile Thr Glu Ala 210 215 220

Pro Asn Gly Tyr Ser Leu Ile Glu Cys Glu Gly Asn Ile Thr Ile Pro

WO 99/63055

| 225 | | | | - | 230 | | | | | 235 | | | | | 240 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Cys | Arg | Leu | Ala | Thr 245 | | Ala | Ser | Gly | Ala 250 | Ala | Ser | Gly | Lys | Phe 255 | Leu |
| Glu | Tyr | Glu | Leu 260 | Gly | Gly | Pro | Arg | Val 265 | Суѕ | Val | Gln | Thr | Ala 270 | Tyr | Gly |
| Ile | Glu | Val 275 | Glu | Val | Glu | Asn | Asn 280 | Pro | Tyr | Asp | Pro | Asp 285 | Leu | Met | Val |
| Phe | Met 290 | Asp | Tyr | Arg | Asp | Phe 295 | Ser | Lys | His | Lys | Pro 300 | Glu | Ser | Leu | Glu. |
| Ala 305 | Lys | Tyr | Pro | Thr | Phe 310 | Leu | Tyr | Val | Met | Ala 315 | Met | Ser | Pro | Thr | Lys 320 |
| Ile | Phe | Phe | Glu | Glu 325 | Thr | Суѕ | Leu | Ala | Ser 330 | Arg | Glu | Ala | Met | Pro 335 | Phe |
| Asn | Leu | Leu | Lys 340 | Ser | Lys | Leu | Met | Ser 345 | Arg | Leu | Lys | Ala | Met 350 | Gly | Ile |
| Arg | Ile | Thr 355 | Arg | Thr | Tyr | Glu | Glu 360 | Glu | Trp | Ser | Tyr | Ile 365 | Pro | Val | Gly |
| Gly | Ser 370 | Leu | Pro | Asn | Thr | Glu 375 | Gln | Lys | Asn | Leu | Ala 380 | Phe | Gly | Ala | Ala |
| Ala 385 | Ser | Met | Val | His | Pro 390 | Ala | Thr | Gly | Tyr | Ser 395 | Val | Val | Arg | Ser | Leu 400 |
| Ser | Glu | Ala | Pro | Asn 405 | Tyr | Ala | Ala | Val | Ile 410 | Ala | Lys | Ile | Leu | Arg 415 | Gln |
| Asp | Gln | Ser | Lys 420 | Glu | Met | Ile | Ser | Leu 425 | Gly | Lys | Tyr | Thr | Asn 430 | Ile | Ser |
| Lys | Gln | Ala 435 | Trp. | Glu | Thr | Leu | Trp 440 | Pro | Leu | Glu | Arg | Lys 445 | Arg | Gln | Arg |
| Ala | Phe 450 | Phe | Leu | Phe | Gly | Leu 455 | Ser | His | Ile | Val | Leu 460 | Met | Asp | Leu | Glu |
| Gly 465 | Thr | Arg | Thr | Phe | Phe 470 | Arg | Thr | Phe | Phe | Arg 475 | Leu | Pro | Lys | Trp | Met 480 |
| Trp | Trp | Gly | Phe | Leu 485 | Gly | Ser | Ser | Leu | Ser 490 | Ser | Thr | Asp | Leu | Ile 495 | Ile |
| Phe | Ala | Leu | Tyr 500 | Met | Phe | Val | Ile | Ala 505 | Pro | His | Ser | Leu | Arg 510 | Met | Glu |
| Leu | Val | Arg 515 | His | Leu | Leu | Ser | Asp 520 | Pro | Thr | Gly | Ala | Thr 525 | Met | Val | Lys |
| Ala | Tyr 530 | Leu | Thr | Ile | | | | | | | | | | | |

<210> 53



<213> Tomato

<40.0> 53

Met Glu Cys Val Gly Val Gln Asn Val Gly Ala Met Ala Val Leu Thr 1 5 10

Arg Pro Arg Leu Asn Arg Trp Ser Gly Glu Leu Cys Gln Glu Lys 20 25 30

Ser Ile Phe Leu Ala Tyr Glu Gln Tyr Glu Ser Lys Cys Asn Ser Ser 35 40 45

Ser Gly Ser Asp Ser Cys Val Val Asp Lys Glu Asp Phe Ala Asp Glu 50 55 60

Glu Asp Tyr Ile Lys Ala Gly Gly Ser Gln Leu Val Phe Val Gln Met 65 70 75 80

Gln Gln Lys Lys Asp Met Asp Gln Gln Ser Lys Leu Ser Asp Glu Leu 85 90 95

Arg Gln Ile Ser Ala Gly Gln Thr Val Leu Asp Leu Val Val Ile Gly
100 105 110

Cys Gly Pro Ala Gly Leu Ala Leu Ala Ala Glu Ser Ala Lys Leu Gly 115 120 125

Leu Asn Val Gly Leu Val Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr 130 135 140

Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Gln Ala Cys Ile 145 150 155 160

Glu His Val Trp Arg Asp Thr Ile Val Tyr Leu Asp Asp Asp Glu Pro 165 170 175

Ile Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Phe Leu His 180 185 190

Glu Glu Leu Leu Lys Arg Cys Val Glu Ala Gly Val Leu Tyr Leu Asn 195 200 205

Ser Lys Val Asp Arg Ile Val Glu Ala Thr Asn Gly Gln Ser Leu Val 210 215 220

Glu Cys Glu Gly Asp Val Val Ile Pro Cys Arg Phe Val Thr Val Ala 225 230 235 240

Ser Gly Ala Ala Ser Gly Lys Phe Leu Gln Tyr Glu Leu Gly Ser Pro 245 250 255

Arg Val Ser Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Asp Asn 260 265 270

Asn Pro Phe Asp Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp Tyr 275 280 285

Leu Arg His Asp Ala Gln Ser Leu Glu Ala Lys Tyr Pro Thr Phe Leu 290 295 300



| Туг 305 | Ala | Met | Pro | Met | Ser 310 | Pro | Thr | Arg | Val | Phe 315 | Phe | Glu | Glu | Thr | Cys 320 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-------------|------------|------------|------------|------------|
| Leu | Ala | Ser | Lys | Asp 325 | Ala | Met | Pro | Phe | Asp 330 | Leu | Leu | Lys | Lys | Lys 335 | Leu |
| Met | Leu | Arg | Leu 340 | Asn | Thr | Leu | Gly | Val 345 | Arg | Ile | Lys | Glu | Ile 350 | Tyr | Glu |
| Glu | Glu | Trp 355 | Ser | Tyr | Ile | Pro | Val 360 | Gly | Gly | Ser | Leu | Pro 365 | Asn | Thr | Glu |
| Gln | Lys 370 | Thr | Leu | Ala | Phe | Gly 375 | Ala | Ala | Ala | Ser | 'Met 380 | Val | His | Pro | Ala |
| Thr 385 | Gly | Tyr | Ser | Val | Val 390 | Arg | Ser | Leu | Ser | Glu 395 | Ala | Pro | Lys | Cys | Ala 400 |
| Ser | Val | Leu | Ala | Asn 405 | Ile | Leu | Arg | Gln | His 410 | Tyr | Ser | Lys | Asn | Met 415 | Leu |
| Thr | Ser | Ser | Ser 420 | Ile | Pro | Ser | Ile | Ser 425 | Thr | Gln | Ala | Trp | Asn 430 | Thr | Leu |
| Trp | Pro | Gln 435 | Glu | Arg | Lys | Arg | Gln 440 | Arg | Ser | Phe | Phe | Leu 445 | Phe | Gly | Leu |
| Ala | Leu 450 | Ile | Leu | Gln | Leu | Asp 455 | Ile | Glu | Gly | Ile | Arg 460 | Ser | Phe | Phe | Arg |
| Ala 465 | Phe | Phe | Arg | Val | Pro 470 | Lys | Trp | Met | Trp | Gln 475 | Gly | Phe | Leu | Gly | Ser 480 |
| Ser | Leu | Ser | Ser | Ala 485 | Asp | Leu | Met | Leu | Phe 490 | Ala | Phe | Tyr | Met | Phe 495 | Ile |
| Ile | Ala | Pro | Asn 500 | Asp | Met | Arg | Lys | Gly 505 | Leu | Ile | Arg | His | Leu 510 | Leu | Ser |
| Asp | Pro | Thr 515 | Gly | Ala | Thr | Leu | Ile 520 | Arg | Thr | Tyr | Leu | Thr 525 | Phe | | |

<210> 54

<211> 516

<212> PRT

<213> Tagetes erecta

<400> 54

Met Ser Met Arg Ala Gly His Met Thr Ala Thr Met Ala Ala Phe Thr 1 5 10 15

Cys Pro Arg Phe Met Thr Ser Ile Arg Tyr Thr Lys Gln Ile Lys Cys 20 25 30

Asn Ala Ala Lys Ser Gln Leu Val Val Lys Gln Glu Ile Glu Glu Glu 35 40 45

Glu Asp Tyr Val Lys Ala Gly Gly Ser Glu Leu Leu Phe Val Gln Met 50 55 60

WO 99/63055



| | | - | | | | | | | | | | | | | |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|-------------|------------|------------|------------|------------|------------|------------|
| Gln 65 | Gln | Asn | Lys | Ser | Met 70 | Asp | Ala | Gln | Ser | Ser 75 | Leu | Ser | Gln | Lys | Leu 80 |
| Pro | Arg | Val | Pro | Ile 85 | Gly | Gly | Gly | Gly | Asp 90 | Ser | Asn | Cys | Ile | Leu 95 | Asp |
| Leu | Val | Val | Ile 100 | Gly | Cys | Gly | Pro | Ala 105 | Gly | Leu | Ala | Leu | Ala 110 | Gly | Glu |
| Ser | Ala | Lys 115 | Leu | Gly | Leu | Asn | Val 120 | Ala | Leu | Ile | Gly | Pro 125 | Asp | Leu | Pro |
| Phe | Thr 130 | Așn | Asn | Tyr | Gly | Val 135 | Trp | Glu | Asp | Glu | Phe 140 | Ile | Gly | Leu | Gly |
| Leu 145 | Glu | Gly | Cys | Ile | Glu 150 | His | Val | Trp | Arg | Asp 155 | Thr | Val | Val | Tyr | Leu 160 |
| Asp | Asp | Asn | Asp | Pro 165 | Ile | Leu | Ile | Gly | 'Arg 170 | Ala | Tyr | Gly | Arg | Val 175 | Ser |
| Arg | Asp | Leu | Leu 180 | His | Glu | Glu | Leu | Leu 185 | Thr | Arg | Cys | Met | Glu 190 | Ser | Gly |
| Val | Ser | Tyr 195 | Leu | Ser | Ser | Lys | Val 200 | Glu | Arg | Ile | Thr | Glu 205 | Ala | Pro | Asn |
| Gly | Leu 210 | Ser | Leu | Ile | Glu | Cys 215 | Glu | Gly | Asn | Ile | Thr 220 | Ile | Pro | Cys | Arg |
| Leu 225 | Ala | Thr | Val | Ala | Ser 230 | Gly | Ala | Ala | Ser | Gly 235 | _ | Leu | Leu | Gln | Tyr 240 |
| Glu | Leu | Gly | Gly | Pro 245 | Arg | Val | Cys | Val | Gln 250 | Thr | Ala | Tyr | Gly | Ile 255 | Glu |
| Val | Glu | Val | Glu 260 | Ser | Ile | Pro | Tyr | Asp 265 | Pro | Ser | Leu | Met | Val 270 | Phe | Met |
| Asp | Tyr | Arg 275 | Asp | Tyr | Thr | Lys | His 280 | Lys | Ser | Gln | Ser | Leu 285 | Glu | Ala | Gln |
| Tyr | Pro 290 | Thr | Phe | Leu | Tyr | Val 295 | | Pro | Met | Ser | Pro 300 | Thr | Lys | Val | Phe |
| Phe 305 | Glu | Glu | Thr | Cys | Leu 310 | Ala | Ser | Lys | Glu | Ala 315 | Met | Pro | Phe | Glu | Leu 320 |
| Leu | Lys | Thr | Lys | Leu 325 | | Ser | Arg | Leu | Lys 330 | | Met | Gly | Ile | Arg 335 | Ile |
| Thr | Lys | Thr | Tyr 340 | | Glu | Glu | Trp | Ser 345 | | Ile | Pro | Val | Gly 350 | Gly | Ser |
| Leu | Pro | Asn 355 | | Glu | Gln | Lys | Asn 360 | | Ala | Phe | Gly | Ala 365 | | Ala | Ser |
| Met | Val 370 | | Pro | Ala | Thr | Gly 375 | | Ser | Val | Val | Arg 380 | | Leu | Ser | Glu |

Ala Pro Asn Tyr Ala Ala Val Ile Ala Lys Ile Leu Gly Lys Gly Asn

400

395

Ser Lys Gln Met Leu Asp His Gly Arg Tyr Thr Thr Asn Ile Ser Lys 405 410 415

390

Gln Ala Trp Glu Thr Leu Trp Pro Leu Glu Arg Lys Arg Gln Arg Ala 420 425 430

Phe Phe Leu Phe Gly Leu Ala Leu Ile Val Gln Met Asp Ile Glu Gly 435 440 445

Thr Arg Thr Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp 450 455 460

Trp Gly Phe Leu Gly Ser Ser Leu Ser Ser Thr Asp Leu Ile Ile Phe 465 470 475 480

Ala Phe Tyr Met Phe Ile Ile Ala Pro His Ser Leu Arg Met Gly Leu 485 490 495

Val Arg His Leu Leu Ser Asp Pro Thr Gly Gly Thr Met Leu Lys Ala 500 505 510

Tyr Leu Thr Ile 515

<210> 55

<211> 501

<212> PRT

<213> Arabidopsis thaliana

<400> 55

Met Asp Thr Leu Leu Lys Thr Pro Asn Lys Leu Asp Phe Phe Ile Pro 1 5 10 15

Gln Phe His Gly Phe Glu Arg Leu Cys Ser Asn Asn Pro Tyr His Ser 20 25 30

Arg Val Arg Leu Gly Val Lys Lys Arg Ala Ile Lys Ile Val Ser Ser 35 40 45

Val Val Ser Gly Ser Ala Ala Leu Leu Asp Leu Val Pro Glu Thr Lys
50 55 60

Lys Glu Asn Leu Asp Phe Glu Leu Pro Leu Tyr Asp Thr Ser Lys Ser 65 70 75 80

Gln Val Val Asp Leu Ala Ile Val Gly Gly Pro Ala Gly Leu Ala 85 90 95

Val Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile Asp 100 105 110

Pro Ser Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val Asp 115 120 125

Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Thr Thr Trp Ser 130 135 140

Gly Ala Val Val Tyr Val Asp Glu Gly Val Lys Lys Asp Leu Ser Arg



| | | - | | | | | | | | | • | | | | |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 145 | | | | | 150 | | | | | 155 | | | | | 160 |
| Pro | Tyŗ | Gly | Arg | Val 165 | Asn | Arg | Lys | Gln | Leu 170 | Lys | Ser | Lys | Met | Leu 175 | Gln |
| Lys | Cys | Ile | Thr 180 | Asn | Gly | Val | Lys | Phe 185 | His | Gln | Ser | Lys | Val 190 | Thr | Asn |
| Val | Val | His 195 | Glu | Glu | Ala | Asn | Ser 200 | Thr | Val | Val | Cys | Ser 205 | Asp | Gly | Val |
| Lys | Ile 210 | Gln | Ala | Ser | Val | Val 215 | Leu | Asp | Ala | Thr | Gly 220 | Phe | Ser | Arg | Cys |
| Leu 225 | Val | Gln | Tyr | Asp | Lys 230 | Pro | Tyr | Asn | Pro | Gly 235 | Tyr | Gln | Val | Ala | Tyr 240 |
| Gly | Ile | Val | Ala | Glu 245 | Val | Asp | Gly | | Pro 250 | Phe | Asp | Val | Asp | Lys 255 | Met |
| Val | Phe | Met | Asp 260 | Trp | Arg | Asp | Lys | His 265 | Leu | Asp | Ser | Tyr | Pro 270 | Glu | Leu |
| Lys | Glu | Arg 275 | Asn | Ser | Lys | Ile | Pro 280 | Thr | Phe | Leu | Tyr | Ala 285 | Met | Pro | Phe |
| Ser | Ser 290 | Asn | Arg | Ile | Phe | Leu 295 | Glu | Glu | Thr | Ser | Leu 300 | Val | Ala | Arg | Pro |
| Gly 305 | Leu | Arg | Met | Glu | Asp 310 | Ile | Gln | Glu | Arg | Met 315 | Ala | Ala | Arg | Leu | Lys 320 |
| His | Leu | Gly | Ile | Asn 325 | Val | Lys | Arg | Ile | Glu 330 | Glű | Asp | Glu | Arg | Cys 335 | Val |
| Ile | Pro | Met | Gly 340 | Gly | Pro | Leu | Pro | Val 345 | Leu | Pro | Gln | Arg | Val 350 | Val | Gly |
| Ile | Gly | Gly 355 | Thr | Ala | Gly | Met | Val 360 | His | Pro | Ser | Thr | Gly 365 | Tyr | Met | Val |
| Ala | Arg 370 | Thr | Leu | Ala | Ala | Ala 375 | Pro | Ile | Val | Ala | Asn 380 | Ala | Ile | Val | Arg |
| Tyr 385 | Leu | Gly | Ser | Pro | Ser 390 | Ser | Asn | Ser | Leu | Arg 395 | Gly | Asp | Gln | Leu | Ser 400 |
| Ala | Glu | Val | Trp | Arg 405 | Asp | Leu | Trp | Pro | Ile 410 | Glu | Arg | Arg | Arg | Gln 415 | Arg |
| Glu | Phe | Phe | Cys 420 | Phe | Gly | Met | Asp | Ile 425 | Leu | Leu | Lys | Leu | Asp 430 | Leu | Asp |
| Ala | Thr | Arg 435 | Arg | Phe | Phe | Asp | Ala 440 | Phe | Phe | Asp | Leu | Gln 445 | Pro | His | Tyr |
| Trp | His 450 | Gly | Phe | Leu | Ser | Ser 455 | Arg | Leu | Phe | Leu | Pro 460 | Glu | Leu | Leu | Val |
| Phe 465 | Gly | Leu | Ser | Leu | Phe 470 | Ser | His | Ala | Ser | Asn 475 | Thr | Ser | Arg | Leu | Glu 480 |





Ile Met Thr Lys Gly Thr Val Pro Leu Ala Lys Met Ile Asn Asn Leu 485 490 495

Val Gln Asp Arg Asp 500

<210> 56

<211> 502

<212> PRT

<213> Adonis palaestina

<400> 56

Met Asp Thr Leu Leu Arg Thr His Asn Lys Leu Glu Leu Leu Pro Thr 1 5 10 15

Leu His Gly Phe Ala Glu Lys Gln His Leu Val Ser Thr Ser Lys Leu
20 25 30

Gln Asn Gln Val Phe Arg Ile Ala Ser Arg Asn Ile His Pro Cys Arg 35 40 45

Asn Gly Thr Val Lys Ala Arg Gly Ser Ala Leu Leu Glu Leu Val Pro 50 55 60

Glu Thr Lys Lys Glu Asn Leu Glu Phe Asp Leu Pro Ala Tyr Asp Pro 65 70 75 80

Ser Arg Gly Ile Val Val Asp Leu Ala Val Val Gly Gly Pro Ala 85 90 95

Gly Leu Ala Ile Ala Gln Gln Val Ser Glu Ala Gly Leu Leu Val Cys 100 105 110

Ser Ile Asp Pro Ser Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val 115 120 125

Trp Val Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Thr 130 135 140

Thr Trp Ser Gly Ala Val Val Tyr Thr Asp Asp Asn Ser Lys Lys Tyr 145 150 155 160

Leu Asp Arg Pro Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys 165 170 175

Met Leu Gln Lys Cys Val Thr Asn Gly Val Lys Phe His Gln Ala Lys 180 185 190

Val Ile Lys Val Ile His Glu Glu Ser Lys Ser Leu Leu Ile Cys Asn 195 200 205

Asp Gly Ile Thr Ile Asn Ala Thr Val Val Leu Asp Ala Thr Gly Phe 210 215 220

Ser Arg Cys Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln 225 230 235 240

Val Ala Tyr Gly Ile Met Ala Glu Val Glu Glu His Pro Phe Asp Leu 245 250 255



Asp Lys Met Leu Phe Met Asp Trp Arg Asp Ser His Leu Asn Glu Lys Leu Glu Leu Lys Asp Lys Asn Arg Lys Ile Pro Thr Phe Leu Tyr Ala Met Pro Phe Ser Ser Thr Lys Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg Pro Gly Leu Arg Phe Glu Asp Ile Gln Glu Arg Met Val Ala Arg Leu Lys His Leu Gly Ile Lys Val Lys Ser Ile Glu Glu Asp Glu Arg Cys Val Ile Pro Met Gly Gly Pro Leu Pro Val Leu Pro Gln Arg Val Val Gly Ile Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly 360 (Tyr Met Val Ala Arg Thr Leu Ala Ala Ala Pro Val Val Ala Lys Ser

Ile Val Gln Tyr Leu Gly Ser Asp Arg Ser Leu Ser Gly Asn Glu Leu 390 395

Ser Ala Glu Val Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Arg Gln

Arg Glu Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu 420 425

Gln Gly Thr Arg Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro His 440

Tyr Trp His Gly Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Leu 450

Phe Phe Gly Leu Ser Leu Phe Ser His Ala Ser Asn Ala Ser Arg Ile 475

Glu Ile Met Ala Lys Gly Thr Val Pro Leu Val Asn Met Met Asn Asn 490

Leu Ile Gln Asp Thr Asp 500

<210> 57

<211> 498

<212> PRT

<213> Pepper

<400> 57

Met Asp Thr Leu Leu Arg Thr Pro Asn Asn Leu Glu Phe Leu His Gly

Phe Gly Val Lys Val Ser Ala Phe Ser Ser Val Lys Ser Gln Lys Phe 25

Glu Asn Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys Gly Val

Val Val Asp Leu Ala Val Val Gly Gly Pro Ala Gly Leu Ala Val

Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile Asp Pro

Asn Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val Asp Glu 115

Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp Ser Gly 135

Ala Ala Val Tyr Ile Asp Asp Lys Thr Thr Lys Asp Leu Asn Arg Pro

Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys Met Met Gln Lys

Cys Ile Leu Asn Gly Val Lys Phe His Gln Ala Lys Val Ile Lys Val 185

Ile His Glu Glu Ser Lys Ser Met Leu Ile Cys Asn Asp Gly Ile Thr

Ile Gln Ala Thr Val Val Leu Asp Ala Thr Gly Phe Ser Arg Ser Leu

Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln Val Ala Tyr Gly 230 235

Ile Leu Ala Glu Val Glu Glu His Pro Phe Asp Val Asn Lys Met Val

Phe Met Asp Trp Arg Asp Ser His Leu Lys Asn Asn Val Glu Leu Lys

Glu Arg Asn Ser Arg Ile Pro Thr Phe Leu Tyr Ala Met Pro Phe Ser 280

Ser Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg Pro Gly

Leu Gly Met Asp Asp Ile Gln Glu Arg Met Val Ala Arg Leu Ser His 310 315

Leu Gly Ile Lys Val Lys Ser Ile Glu Glu Asp Glu His Cys Val Ile 325

Pro Met Gly Gly Pro Leu Pro Val Leu Pro Gln Arg Val Val Gly Ile 345

Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met Val Ala

BNSDOCID: <WO

_9963055A1_I_>

355 360 365

Arg Thr Leu Ala Ala Ala Pro Val Val Ala Asn Ala Ile Ile Gln Tyr 370 380

Leu Ser Ser Glu Arg Ser His Ser Gly Asp Glu Leu Ser Ala Ala Val 385 390 395 400

Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Gln Arg Glu Phe Phe
405 410 415

Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu Pro Ala Thr Arg
420 425 430

Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp His Gly 435 440 445

Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Ile Val Phe Gly Leu 450 460

Ser Leu Phe Ser His Ala Ser Asn Thr Ser Arg Leu Glu Ile Met Thr 465 470 475 480

Lys Gly Thr Leu Pro Leu Val His Met Ile Asn Asn Leu Leu Gln Asp 485 490 495

Lys Glu

<210> 58

<211> 500

<212> PRT

<213> Tomato

<400> 58

Met Asp Thr Leu Leu Lys Thr Pro Asn Asn Leu Glu Phe Leu Asn Pro
1 5 10 15

His His Gly Phe Ala Val Lys Ala Ser Thr Phe Arg Ser Glu Lys His 20 25 30

His Asn Phe Gly Ser Arg Lys Phe Cys Glu Thr Leu Gly Arg Ser Val 35 40 45

Cys Val Lys Gly Ser Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr 50 55 60

Lys Lys Glu Asn Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys 65 70 75 80

Gly Val Val Val Asp Leu Ala Val Val Gly Gly Pro Ala Gly Leu 85 90 95

Ala Val Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile 100 105

Asp Pro Asn Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val 115 120 125

Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp



| | 130 | | | | | 135 | | | | | 140 | | | | + ∵ |
|------------|------------|------------|------------|------------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Ser 145 | Gly | Ala | Ala | Val | Tyr .150 | Ile | Asp | Asp | Asn | Thr 155 | Ala | Lys | Asp | Leu | His 160 |
| Arg | Pro | Tyr | Gly | Arg 165 | Val | Asn | Arg | Lys | Gln 170 | Leu | Lys | Ser | Lys | Met 175 | Met |
| Gln | Lys | Cys | Ile 180 | Met | Asn | Gly | Val | Lys 185 | Phe | His | Gln | Ala | Lys 190 | Val | Ile |
| Lys | Val | Ile 195 | His | Glu | Glu | Ser | Lys 200 | Ser | Met | Leu | Ile | Cys 205 | Asn | Asp | Gly |
| Ile | Thr 210 | Ile | Gln | Ala | Thr | Val 215 | Val | Leu | Asp | Ala | Thr 220 | Gly | Phe | Ser | Arg |
| Ser 225 | Leu | Val | Gln | Tyr | Asp 230 | Lys | Pro | Tyr | Asn | Pro 235 | Gly | Tyr | Gln | Val | Ala 240 |
| Tyr | Gly | Ile | Leu | Ala 245 | Glu | Val | Glu | Glu | His 250 | Pro | Phe | Asp | Val | Asn 255 | Lys |
| Met | Val | Phe | Met 260 | Asp | Trp | Arg | Asp | Ser 265 | His | Leu | Lys | Asn | Asn 270 | Thr | Asp |
| Leu | Lys | Glu 275 | Arg | Asn | Ser | Arg | Ile 280 | Pro | Thr | Phe | Leu | Tyr 285 | Ala | Met | Pro |
| Phe | Ser 290 | Ser | Asn | Arg | Ile | Phe 295 | Leu | Glu | Glu | Thr | Ser 300 | Leu | Val | Ala | Arg |
| Pro 305 | Gly | Leu | Arg | Ile | Asp 310 | Asp | Ile | Gln | Glu | Arg 315 | Met | Val | Ala | Arg | Leu 320 |
| Asn | His | Leu | Gly | 11e 325 | Lys | Val | Lys | Ser | 11e 330 | Glu | Glu | Asp | Glu | His 335 | Cys |
| Leu | Ile | Pro | Met 340 | Gly | Gly | Pro | Leu | Pro 345 | Val | Leu | Pro | Gln | Arg 350 | Val | Val |
| Gly | Ile | Gly 355 | Gly | Thr | Ala | Gly | Met 360 | Val | His | Р́го | Ser | Thr 365 | Gly | Tyr | Met |
| Val | Ala 370 | Arg | Thr | Leu | Ala | Ala 375 | Ala | Pro | Val | Val | Ala 380 | Asn | Ala | Ile | Ile |
| Gln 385 | Tyr | Leu | Gly | Ser | Glu 390 | Årg | Ser | His | Ser | Gly 395 | Asn | Glu | Leu | Ser | Thr 400 |
| Ala | Val | Trp | Lys | Asp 405 | Leu | Trp | Pro | Ile | Glu 410 | Arg | Arg | Arg | Gln | Arg 415 | Glu |
| Phe | Phe | Cys | Phe 420 | Gly | Met | Asp | Ile | Leu 425 | Leu | Lys | Leu | Asp | Leu 430 | Pro | Ala |
| Thr | Arg | Arg 435 | Phe | Phe | Asp | Ala | Phe 440 | Phe | Asp | Leu | Glu | Pro 445 | Arg | Tyr | Trp |
| His | Gly 450 | Phe | Leu | Ser | Ser | Arg 455 | | Phe | Leu | Pro | Glu 460 | Leu | Ile | Val | Phe |
| | | | | | | | | | | | | | | | |



Gly Leu Ser Leu Phe Ser His Ala Ser Asn Thr Ser Arg Phe Glu Ile 465 470 475 480

Met Thr Lys Gly Thr Val Pro Leu Val Asn Met Ile Asn Asn Leu Leu 485 490 495

Gln Asp Lys Glu 500

<210> 59

<211> 500

<212> PRT

<213> Tobacco

<400> 59

Met Asp Thr Leu Leu Lys Thr Pro Asn Lys Leu Glu Phe Leu His Pro 1 5 10 15

Val His Gly Phe Ser Val Lys Ala Ser Ser Phe Asn Ser Val Lys Pro 20 25 30

His Lys Phe Gly Ser Arg Lys Ile Cys Glu Asn Trp Gly Lys Gly Val 35 40 45

Cys Val Lys Ala Lys Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr 50 55 60

Lys Lys Glu Asn Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys 65 70 75 80

Gly Leu Val Val Asp Leu Ala Val Val Gly Gly Pro Ala Gly Leu 85 90 95

Ala Val Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Val Ser Ile 100 105 110

Asp Pro Ser Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val 115 120 125

Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp 130 140

Ser Gly Thr Val Val Tyr Ile Asp Asp Asn Thr Thr Lys Asp Leu Asp 145 150 155 160

Arg Pro Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys Met Met 165 170 175

Gln Lys Cys Ile Leu Asn Gly Val Lys Phe His His Ala Lys Val Ile 180 185 190

Lys Val Ile His Glu Glu Ala Lys Ser Met Leu Ile Cys Asn Asp Gly 195 200 205

Val Thr Ile Gln Ala Thr Val Val Leu Asp Ala Thr Gly Phe Ser Arg 210 215 220

Cys Leu Val Gln Tyr Asp Lys Pro Tyr Lys Pro Gly Tyr Gln Val Ala 225 230 235 240



| Tyr | Glv | Ile | Len | Ala | Glu | ۷a۱ | Glu | Glu | Hic | D~~ | Dha | 7\ | m '- | | Lys |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| | | | | 245 | | | | | 250 | | | | | 255 | • |
| | | | Met 260 | | | | | 265 | | | | | 270 | | |
| Leu | Lys | Glu 275 | Arg | Asn | Arg | Lys | Val 280 | Pro | Thr | Phe | Leu | Tyr 285 | Ala | Met | Pro |
| Phe | Ser 290 | Ser | Asn | Lys | Ile | Phe 295 | Leu | Glu | Glu | Thr | Ser 300 | Leu | Val | Ala | Arg |
| Pro 305 | Gly | Leu | Arg | Met | Asp 310 | Asp | Ile | Gln | Glu | Arg 315 | Met | Val | Ala | Arg | Leu 320 |
| Asn | His | Leu | Gly | Ile 325 | Lys | Val | Lys | Ser | Ile 330 | Glu | Glu | Asp | Glu | His 335 | Cys |
| Val | Ile | Pro | Met 340 | Gly | Gly | Ser | Leu | Pro 345 | Val | Ile | Pro | Gln | Arg 350 | Val | Val |
| Gly | Thr | Gly 355 | Gly | Thr | Ala | Gly | Leu 360 | Val | His | Pro | Ser | Thr 365 | Gly | Tyr | Met |
| Val | Ala 370 | Arg | Thr | Leu | Ala | Ala 375 | Ala | Pro | Val | Val | Ala 380 | Asn | Ala | Ile | Ile |
| His 385 | Tyr | Leu | Gly | Ser | Glu 390 | Lys | Asp | Leu | Leu | Gly 395 | Asn | Glu | Leu | Ser | Ala 400 |
| Ala | Val | Trp | Lys | Asp 405 | Leu | Trp | Pro | Ile | Glu 410 | Arg | Arg | Arg | Gln | Arg 415 | |
| Phe | Phe | Cys | Phe 420 | Gly | Met | Asp | Ile | Leu 425 | Leu | Lys | Leu | Asp | Leu 430 | Pro | Ala |
| Thr | Arg | Arg 435 | Phe | Phe | Asp | Ala | Phe 440 | Phe | Asp | Leu | Glu | Pro 445 | Arg | Tyr | Trp |
| His | Gly 450 | Phe | Leu | Ser | Ser | Arg 455 | Leu | Tyr | Leu | Pro | Glu 460 | Leu | Ile | Phe | Phe |
| Gly 465 | Leu | Ser | Leu | Phe | Ser 470 | Arg | Ala | Ser | Asn | Thr 475 | Ser | Arg | Ile | Glu | Ile 480 |
| Met | Thr | Lys | Gly | Thr 485 | Leu | Pro | Leu | Val | Asn 490 | Met | Ile | Asn | Asn | Leu 495 | Leu |
| Gln | Asp | Thr | Glu | | | | | | | | | | | | |

Gln Asp Thr Glu 500

<210> 60 <211> 511 <212> PRT <213> Tagetes erecta

<400> 60

Met Asp Thr Phe Leu Arg Thr Tyr Asn Ser Phe Glu Phe Val His Pro 1 10



| | | | | ,,,,,, | -/ | | | | | | | | | — . | |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Ser | Asn | Lys | Phe 20 | Àla | Gly | Asn | Leu | Asn 25 | Asn | Leu | Asn | Gln | Leu 30 | Asn | ĞÎn |
| Ser | Lys | Ser .35 | Gln | Phe | Gln | Asp | Phe 40 | .Arg | Phe | Gly | Pro | Lys 45 | Lys | Ser | Gln |
| Phe | Lys 50 | Leu | Gly | Gln | Lys | Tyr 55 | Cys | Val | Lys | Ala | Ser 60 | Ser | Ser | Ala | Leu |
| Leu 65 | Glu | Leu | Val | Pro | Glu 70 | Ile | Lys | Lys | Glu | Asn 75 | Leu | Asp | Phe | Asp | Leu 80 |
| Pro | Met | Tyr | Asp | Pro 85 | Ser | Arg | Asn | Val | Val 90 | Val | Asp | Leu | Val | Val 95 | Val |
| Gly | Gly | Gly | Pro 100 | Ser | Gly | Leu | Ala | Val 105 | Ala | Gln | Gln | Val | Ser 110 | Glu | Ala |
| Glý | Leu | Thr 115 | Val | Cys | Ser | Ile | Asp 120 | Pro | Ser | Pro | Lys | Leu 125 | Ile | Trp | Pro |
| Asn | Asn 130 | Tyr | Gly | Val | Trp | Val 135 | Asp | Glu | Phe | Glu | Ala 140 | Met | Asp | Leu | Leu |
| Asp 145 | Cys | Leu | Asp | Thr | Thr 150 | Trp | Ser | Ser | Ala | Val 155 | Val | Tyr | Ile | Asp | Glu 160 |
| Lys | Ser | Thr | Lys | Ser 165 | Leu | Asn | Arg | Pro | Tyr 170 | Ala | Arg | Val | Asn | Arg 175 | Lys |
| Gln | Leu | Lys | Thr 180 | Lys | Met | Leu | Gln | Lys 185 | Cys | Ile | Ala | Asn | Gly 190 | Val | Lys |
| Phe | His | Gln 195 | Ala | Lys | Val | Ile | Lys 200 | Val | Ile | His | Glu | Glu 205 | Leu | Lys | Ser |
| Leu | Leu 210 | Ile | Cys | Asn | Asp | Gly 215 | Val | Thr | Ile | Gln | Ala 220 | Thr | Leu | Val | Leu |
| Asp 225 | Ala | Thr | Gly | Phe | Ser 230 | Arg | ·Ser | Leu | Val | Gln 235 | Tyr | Asp | Lys | Pro | Tyr 240 |
| Asn | Pro | Gly | Tyr | Gln 245 | Val | Ala | Tyr | Gly | Ile 250 | Leu | Ala | Glu | Val | Glu 255 | Glu |
| His | Pro | Phe | Asp 260 | Val | Asp | Lys | Met | Leu 265 | Phe | Met | Asp | Trp | Arg 270 | Asp | Ser |
| His | Leu | Asp 275 | Gln | Asn | Leu | Glu | Ile 280 | Lys | Ala | Arg | Asn | Ser 285 | Arg | Ile | Pro |
| Thr | Phe 290 | Leu | Tyr | Ala | Met | Pro 295 | Phe | Ser | Ser | Thr | Arg 300 | Ile | Phe | Leu | Glu |
| Glu 305 | Thr | Ser | Leu | Val | Ala 310 | Arg | Pro | Gly | Leu | Lys 315 | Met | Glu | Asp | Ile | Gln 320 |
| Glu | Arg | Met | Ala | Tyr 325 | Arg | Leu | Lys | His | Leu 330 | Gly | Ile · | Lys | Val | Lys 335 | |
| Ile | Glu | Glu | Asp | Glu | Arg | Cys | Val | Ile | Pro | Met | Gly | Gly | Pro | Leu | Pro |



His Pro Ser Thr Gly Tyr Met Val Ala Arg Thr Leu Ala Ala Pro 370 375 380

Ile Val Ala Lys Ser Ile Ile Arg Tyr Leu Asn Asn Glu Lys Ser Met 385 390 395 400

Val Ala Asp Val Thr Gly Asp Asp Leu Ala Ala Gly Ile Trp Arg Glu . 405 410 415

Leu Trp Pro Ile Glu Arg Arg Gln Arg Glu Phe Phe Cys Phe Gly
420 425 430

Met Asp Ile Leu Leu Lys Leu Asp Leu Glu Gly Thr Arg Arg Phe Phe 435 440 445

Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp His Gly Phe Leu Ser 450 460

Ser Arg Leu Phe Leu Pro Glu Leu Val Thr Phe Gly Leu Ser Leu Phe 465 470 475 480

Gly His Ala Ser Asn Thr Cys Arg Val Glu Ile Met Ala Lys Gly Thr 485 490 495

Leu Pro Leu Ala Thr Met Ile Gly Asn Leu Val Arg Asp Arg Glu 500 505 510

<210> 61

<211> 503

<212> PRT

<213> Daffodil

<400> 61

Met Asp Thr Leu Leu Arg Thr His Asn Arg Leu Glu Leu Leu Tyr Pro

Leu His Glu Leu Ala Lys Arg His Phe Leu Ser Pro Ser Pro Asn Pro 20 25 30

Gln Asn Pro Asn Phe Lys Phe Phe Ser Arg Lys Pro Tyr Gln Lys Lys
35 40 45

Cys Arg Asn Gly Tyr Ile Gly Val Ser Ser Asn Gln Leu Leu Asp Leu
50 60

Val Pro Glu Ile Lys Lys Glu His Leu Glu Phe Asp Leu Pro Leu Tyr 65 70 75 80

Asp Pro Ser Lys Ala Leu Thr Leu Asp Leu Ala Val Val Gly Gly Gly 85 90 95

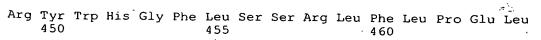
Pro Leu Ala Arg Ser Cys Ser Thr Ser Leu Gly Gly Gly Leu Ser Val

Val Ser Ile Asp Pro Asn Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly

بن في 115 120 125 Val Trp Val Asp Glu Phe Glu Asp Met Asp Leu Leu Asp Cys Leu Asp 135 Ala Thr Trp Ser Gly Ala Ile Val Tyr Val Asp Asp Arg Ser Thr Lys 150 Asn Leu Ser Arg Pro Tyr Ala Arg Val Asn Arg Lys Asn Leu Lys Ser Lys Met Met Lys Lys Cys Val Ser Asn Gly Val Arg Phe His Gln Ala Thr Val Val Lys Ala Met His Glu Glu Glu Lys Ser Tyr Leu Ile Cys 200 Ser Asp Gly Val Thr Ile Asp Ala Arg Val Val Leu Asp Ala Thr Gly Phe Ser Arg Cys Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln Val Ala Tyr Gly Ile Leu Ala Glu Val Glu Glu His Pro Phe Asp Val Asp Lys Met Val Phe Met Asp Trp Arg Asp Ser His Leu Asn Gly Lys Ala Glu Leu Asn Glu Arg Asn Ala Lys Ile Pro Thr Phe Leu Tyr 280 Ala Met Pro Phe Ser Ser Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu 290 Val Ala Arg Pro Gly Leu Lys Met Glu Asp Ile Gln Glu Arg Met Val Ala Arg Leu Asn His Leu Gly Ile Arg Ile Lys Ser Ile Glu Glu Asp Glu Arg Cys Val Ile Pro Met Gly Gly Pro Leu Pro Val Ile Pro Gln Arg Val Val Gly Ile Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met Val Ala Arg Thr Leu Ala Ala Ala Pro Ile Val Ala Asn Ser Ile Val Gln Tyr Leu Val Ser Asp Ser Gly Leu Ser Gly Asn Asp 395 Leu Ser Ala Asp Val Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg 410 Gln Arg Glu Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp 425

Leu Glu Gly Thr Arg Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro





Val Pro Phe Gly Leu Ser Leu Phe Ser His Ala Ser Asn Thr Cys Lys 465 470 475 480

Leu Glu Ile Met Ala Lys Gly Thr Leu Pro Leu Val Asn Met Ile Asn 485 490 495

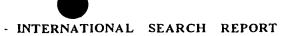
Asn Leu Val Gln Asp Arg Asp 500

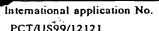


INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/12121

| | | | | | • |
|-------------|---|-------------------|-----------|--|---|
| li . | SSIFICATION OF SUBJECT MATTER | 4 | | | |
| | :Please See Extra Sheet. :435/189, 193, 233, 252.3, 320.1, 325; 536/23.2 | | | | • |
| According t | to International Patent Classification (IPC) or to both | national classifi | ication . | and IPC | |
| B. FIEL | DS SEARCHED | | | | |
| Minimum d | ocumentation searched (classification system followed | by classificati | on sym | bols) | |
| U.S. : | 435/189, 193, 233, 252.3, 320.1, 325; 536/23.2 | | • | · | |
| Documentat | tion searched other than minimum documentation to the | extent that suc | h docum | nents are included | in the fields searched |
| | | | | | |
| | lata base consulted during the international search (na e Extra Sheet. | me of data bas | e and, | where practicable, | , search terms used) |
| C. DOC | UMENTS CONSIDERED TO BE RELEVANT | | | | |
| Category* | Citation of document, with indication, where ap | propriate, of the | e releva | int passages | Relevant to claim No. |
| X | WO 97/36998 A1 (UNIVERSITY O PARK) 09 October 1997, see entire d No:1. | | | | 1-8 |
| | | | | | |
| | · | | | | |
| | | | | | |
| | | | | , | |
| | | | | • | |
| | | | | | |
| | | | | | • |
| | | | | | |
| | | | | | |
| Furth | ner documents are listed in the continuation of Box C | . Se | e paten | t family annex. | |
| ! | ecial categories of cited documents: | | | | ernational filing date or priority |
| | cument defining the general state of the art which is not considered be of particular relevance | | | theory underlying the | |
| "E. ean | rlier document published on or after the international filing date | | | | e claimed invention cannot be red to involve an inventive step |
| | cument which may throw doubts on priority claim(s) or which is ed to establish the publication date of another citation or other | when | | ment is taken slone | |
| spe | ecial reason (as specified) | | | | e claimed invention cannot be step when the document is |
| | ecument referring to an oral disclosure, use, exhibition or other eans | | | one or more other such to a person skilled in t | n documents, such combination the art |
| | cument published prior to the international filing date but later than e priority date claimed | .w. qocan | nent men | ber of the same patent | t family |
| Date of the | actual completion of the international search | Date of mailin | g of th | e international sea | irch report |
| 02 AUGU | UST 1999 | 15 5 | EP | 1999 | |
| | mailing address of the ISA.US | Authorized of | licer | | |
| Box PCT | ner of Patents and Trademarks n. D.C. 2023 i | BRADLE | Y S. M. | AYHEW | |
| Facsimile N | | Telephone No | . (7 | 03) 308-0196 | YU C |





| | PCT/US99/12121 |
|--|--|
| A. CLASSIFICATION OF SUBJECT MATTER: IPC (6): | - |
| C12N 1/21, 5/10, 9/10, 15/53, 15/54, 15/61, 15/63; C12P 23/00; C12Q 1/68 | |
| B. FIELDS SEARCHED Electronic data bases consulted (Name of data base and where practicable term | os used): |
| Dialog and APS search terms: IPP, epsilon cyclase, lycopene cyclase, isopentenyl pyrophospha isomerase | te isomerase and isopentenyl diphosphate |
| | |
| | · |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| · | |
| | |
| | |
| | |
| | |
| · | |
| | |
| | |
| | |

Form PCT:ISA 210 (extra sheet)(July 1992)*

PCT





INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6:

C12N 1/21, 5/10, 9/10, 15/53, 15/54, 15/61, 15/63, C12P 23/00, C12Q 1/68 (11) International Publication Number:

WO 99/63055

(43) International Publication Date: 9 December 1999 (09.12.99)

(21) International Application Number:

PCT/US99/12121

A1

(22) International Filing Date:

2 June 1999 (02.06.99)

(30) Priority Data:

09/088,724 09/088,725

2 June 1998 (02.06.98) 2 June 1998 (02.06.98) 211 US

(71) Applicant (for all designated States except US): UNIVERSITY OF MARYLAND [US/US]; Office of Technology Liaison, 4312 Knox Road, College Park, MD 20742 (US).

(72) Inventors; and

- (75) Inventors/Applicants (for US only): CUNNINGHAM, Francis, X., Jr. [US/US]; 2727 Washington Avenue, Chevy Chase, MD 20815 (US). SUN, Zairen [US/US]; 3405 Tulane Drive #22, Hyattsville, MD 20783 (US).
- (74) Agents: GOLDHUSH, Douglas, H. et al.; Nikaido, Marmelstein, Murray & Oram LLP, Suite 330 G Street Lobby, Metropolitan Square, 655 Fifteenth Street, N.W., Washington, DC 20005-5701 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report. With amended claims.

(54) Title: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND METHODS OF USE THEREOF

(57) Abstract

Nucleic acid sequences encoding ϵ -cyclase, isopentenyl pyrophosphate isomerase and β -carotene hydroxylase as well as vectors containing the same and hosts transformed with the vectors. Methods for controlling the ratio of various carotenoids in a host and for the production of novel carotenoid pigments. The present invention also provides a method for screening for eukaryotic genes encoding carotenoid biosynthesis, and for modifying the disclosed enzymes.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| AL | Albania | ES | Spain | LS | Lesotho | SI | Slovenia |
|----|--------------------------|----|---------------------|----|-----------------------|----|--------------------------|
| AM | Armenia | FI | Finland | LT | Lithuania | SK | Slovakia |
| AT | Austria | FR | France | LU | Luxembourg | SN | Senegal |
| AU | Australia | GA | Gabon | LV | Latvia | SZ | Swaziland |
| AZ | Azerbaijan | GB | United Kingdom | MC | Monaco | TD | Chad |
| BA | Bosnia and Herzegovina | GE | Georgia | MD | Republic of Moldova | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagascar | TJ | Tajikistan |
| BE | Belgium | GN | Guinea | MK | The former Yugoslav | TM | Turkmenistan |
| BF | Burkina Faso | GR | Greece | | Republic of Macedonia | TR | Turkey |
| BG | Bulgaria | HU | Hungary | ML | Mali | TT | Trinidad and Tobago |
| BJ | Benin | IE | Ireland | MN | Mongolia | UA | Ukraine |
| BR | Brazil | IL | Israel | MR | Mauritania | UG | Uganda |
| BY | Belarus | IS | Iceland | MW | Malawi | US | United States of America |
| CA | Canada | IT | Italy | MX | Mexico | UZ | Uzbekistan |
| CF | Central African Republic | JP | Japan | NE | Niger | VN | Viet Nam |
| CG | Congo | KE | Kenya | NL | Netherlands | YU | Yugoslavia |
| CH | Switzerland | KG | Kyrgyzstan | NO | Norway | zw | Zimbabwe |
| CI | Côte d'Ivoire | KP | Democratic People's | NZ | New Zealand | | |
| СМ | Cameroon | | Republic of Korea | PL | Poland | | |
| CN | China | KR | Republic of Korea | PT | Portugal | | |
| CU | Cuba | KZ | Kazakstan | RO | Romania | | |
| CZ | Czech Republic | LC | Saint Lucia | RU | Russian Federation | | |
| DE | Germany | LI | Liechtenstein | SD | Sudan | | |
| DK | Denmark | LK | Sri Lanka | SE | Sweden | | |
| EE | Estonia | LR | Liberia | SG | Singapore | | |

WO 99/63055 PCT/US99/12121

GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND METHODS OF USE THEREOF

BACKGROUND OF THE INVENTION

Field of the Invention

5

10

The present invention describes nucleic acid sequences for eukaryotic genes encoding ϵ lycopene ϵ -cyclase (also known as ϵ -cyclase and ϵ lycopene cyclase), isopentenyl pyrophosphate isomerase (IPP) and β -carotene hydroxylase as well as vectors containing the same and hosts transformed with said vectors. The present invention also provides methods for augmenting the accumulation of carotenoids, changing the composition of the carotenoids, and producing novel and rare carotenoids. The present invention provides methods for controlling the ratio or relative amounts of various carotenoids in a host. The invention also relates to modified lycopene ϵ -cyclase, IPP isomerase and β -carotene hydroxylase. Additionally, the present invention provides a method for screening for genes and cDNAs encoding enzymes of carotenoid biosynthesis and metabolism.

15

20

25

30

Background of the Invention

Carotenoid pigments with cyclic endgroups are essential components of the photosynthetic apparatus in oxygenic photosynthetic organisms (e.g., cyanobacteria, algae and plants; Goodwin, 1980). The symmetrical bicyclic yellow carotenoid pigment βcarotene (or, in rare cases, the asymmetrical bicyclic α-carotene) is intimately associated with the photosynthetic reaction centers and plays a vital role in protecting against potentially lethal photooxidative damage (Koyama, 1991). β-carotene and other carotenoids derived from it or from α-carotene also serve as light-harvesting pigments (Siefermann-Harms, 1987), are involved in the thermal dissipation of excess light energy captured by the lightharvesting antenna (Demmig-Adams & Adams, 1992), provide substrate for the biosynthesis of the plant growth regulator abscisic acid (Rock & Zeevaart, 1991; Parry & Horgan, 1991), and are precursors of vitamin A in human and animal diets (Krinsky, 1987). Plants also exploit carotenoids as coloring agents in flowers and fruits to attract pollinators and agents of seed dispersal (Goodwin, 1980). The color provided by carotenoids is also of agronomic value in a number of important crops. Carotenoids are currently harvested from a variety of organisms, including plants, algae, yeasts, cyanobacteria and bacteria, for use as pigments in food and feed.

The probable pathway for formation of cyclic carotenoids in plants, algae and cyanobacteria is illustrated in Figure 1. Two types of cyclic endgroups or rings are commonly found in higher plant carotenoids, these are referred to as the β (beta) and ϵ (epsilon) rings (Fig. 3). The precursor acyclic endgroup (no ring structure) is referred to as the Ψ (psi) endgroup. The β and ϵ endgroups differ only in the position of the double bond in the ring. Carotenoids with two β rings are ubiquitous, and those with one β and one ϵ ring are common, but carotenoids with two ϵ rings are uncommon. β -carotene (Fig. 1) has two β -endgroups and is a symmetrical compound that is the precursor of a number of other important plant carotenoids such as zeaxanthin and violaxanthin (Fig. 2).

10

15

20

25

5

Genes encoding enzymes of carotenoid biosynthesis have previously been isolated from a variety of sources including bacteria (Armstrong et al., 1989, Mol. Gen. Genet. 216, 254-268; Misawa et al., 1990, J. Bacteriol., 172, 6704-12), fungi (Schmidhauser et al., 1990, Mol. Cell. Biol. 10, 5064-70), cyanobacteria (Chamovitz et al., 1990, Z. Naturforsch, 45c, 482-86; Cunningham et al., 1994) and higher plants (Bartley et al., Proc. Natl. Acad. Sci USA 88, 6532-36; Martinez-Ferez & Vioque, 1992, Plant Mol. Biol. 18, 981-83). Many of the isolated enzymes show a great diversity in structure, function and inhibitory properties between sources. For example, phytoene desaturases from the cyanobacterium Synechococcus and from higher plants and green algae carry out a two-step desaturation to yield ζ -carotene as a reaction product. In plants and cyanobacteria a second enzyme (ζ carotene desaturase), similar in amino acid sequence to the phytoene desaturase, catalyzes two additional desaturations to yield lycopene. In contrast, a single desaturase enzyme from Erwinia herbicola and from other bacteria introduces all four double bonds required to form lycopene. The Erwinia and other bacterial desaturases bear little amino acid sequence similarity to the plant and cyanobacterial desaturase enzymes, and are thought to be of unrelated ancestry. Therefore, even with a gene in hand from one source, it may be difficult to identify a gene encoding an enzyme of similar function in another organism. In particular, the sequence similarity between certain of the prokaryotic and eukaryotic genes encoding enzymes of carotenoid biosynthesis is quite low.

30

Further, the mechanism of gene expression in prokaryotes and eukaryotes appears to differ sufficiently such that one cannot expect that an isolated eukaryotic gene will be properly expressed in a prokaryotic host.

The difficulties in isolating genes encoding enzymes with similar functions is exemplified by recent efforts to isolate the gene encoding the enzyme that catalyzes the formation of β -carotene from the acyclic precursor lycopene. Although a gene encoding an enzyme with this function had been isolated from a bacterium, it had not been isolated from any photosynthetic procaryote or from any eukaryotic organism. The isolation and characterization of the enzyme catalyzing formation of β -carotene in the cyanobacterium Synechococcus PCC7942 was described by the present inventors and others (Cunningham et al., 1993 and 1994). The amino acid sequence similarity of the cyanobacterial enzyme to the various bacterial lycopene β -cyclases is so low (ca. 18-25% overall; Cunningham et al., 1994) that there is much uncertainty as to whether they share a common ancestry or, instead, represent an example of convergent evolution.

The need remains for the isolation of eukaryotic and prokaryotic genes and cDNAs encoding polypeptides involved in the carotenoid biosynthetic pathway, including those encoding a lycopene ϵ -cyclase, IPP isomerase and β -carotene hydroxylase. There remains a need for methods to enhance the production of carotenoids, to alter the composition of carotenoids, and to reduce or eliminate carotenoid production. There also remains a need in the art for methods for screening for genes and cDNAs encoding enzymes of carotenoid biosynthesis and metabolism.

SUMMARY OF THE INVENTION

20

5

10

15

Accordingly, a first object of this invention is to provide purified and/or isolated nucleic acids which encode enzymes involved in carotenoid biosynthesis; in particular, lycopene ϵ -cyclase, IPP isomerase and β -carotene hydroxylase.

A second object of this invention is to provide purified and/or isolated nucleic acids which encode enzymes which produce novel or uncommon carotenoids.

25

A third object of the present invention is to provide vectors containing said genes.

A fourth object of the present invention is to provide hosts transformed with said vectors.

Another object of the present invention is to provide hosts which accumulate novel or uncommon carotenoids or which accumulate greater amounts of specific or total carotenoids.

Another object of the present invention is to provide hosts with inhibited and/or altered carotenoid production.

PCT/US99/12121

WO 99/63055

5

10

15

20

25

30

Another object of this invention is to secure the expression of eukaryotic carotenoid-related genes in a recombinant prokaryotic host.

Yet another object of the present invention is to provide a method for screening for eukaryotic and prokaryotic genes and cDNAs which encode enzymes involved in carotenoid biosynthesis and metabolism.

An additional object of the invention is to provide a method for manipulating carotenoid biosynthesis in photosynthetic organisms by inhibiting the synthesis of certain enzymatic products to cause accumulation of precursor compounds.

Another object of the invention is to provide modified lycopene ϵ -cyclase, IPP isomerase and β -carotene hydroxylase.

These and other objects of the present invention have been realized by the present inventors as described below.

A subject of the present invention is an isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity and having the amino acid sequence of SEQ ID NOS: 2, 4, 14-21 or 23-27.

The invention also includes vectors which comprise any of the nucleic acid sequences listed above, and host cells transformed with such vectors.

Another subject of the present invention is a method of producing or enhancing the production of a carotenoid in a host cell, comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence to produce the protein.

Yet another subject of the present invention is a method of modifying the production of carotenoids in a host cell, the method comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which produces an RNA and/or encodes for a protein which modifies lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity, relative to an untransformed host cell, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence in the host cell to modify the production of the carotenoids in the host cell, relative to the untransformed host cell.

10

15

20

25

30

The present invention also includes a method of expressing, in a host cell, a heterologous nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity, the method comprising inserting into the host cell a vector comprising the heterologous nucleic acid sequence, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence.

Also included is a method of expressing, in a host cell, a heterologous nucleic acid sequence which encodes for a protein which modifies lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity in the host cell, relative to an untransformed host cell, the method comprising inserting into the host cell a vector comprising the heterologous nucleic acid sequence, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence.

Another subject of the present invention is a method for screening for genes and cDNAs which encode enzymes involved in carotenoid biosynthesis and metabolism.

BRIEF DESCRIPTION OF THE DRAWINGS

A more complete appreciation of the invention and many of the attendant advantages thereof will be readily obtained as the same becomes better understood by reference to the following detailed description when considered in connection with the accompanying drawings, wherein:

Figure 1 is a schematic representation of the putative pathway of β -carotene biosynthesis in cyanobacteria, algae and plants. The enzymes catalyzing various steps are indicated at the left. Target sites of the bleaching herbicides NFZ and MPTA are also indicated at the left. Abbreviations: DMAPP, dimethylallyl pyrophosphate; FPP, farnesyl pyrophosphate; GGPP, geranylgeranyl pyrophosphate; GPP, geranyl pyrophosphate; IPP, isopentenyl pyrophosphate; LCY, lycopene cyclase; MVA, mevalonic acid; MPTA, 2-(4-methylphenoxy)triethylamine hydrochloride; NFZ, norflurazon; PDS, phytoene desaturase; PSY, phytoene synthase; ZDS, ζ -carotene desaturase; PPPP, prephytoene pyrophosphate.

Figure 2 depicts possible routes of synthesis of cyclic carotenoids and common plant and algal xanthophylls (oxycarotenolds) from neurosporene. Demonstrated activities of the β - and ϵ -cyclase enzymes of A. thaliana are indicated by bold arrows labelled with β or ϵ respectively. A bar below the arrow leading to ϵ -carotene indicates that the enzymatic

WO 99/63055 PCT/US99/12121

activity was examined but no product was detected. The steps marked by an arrow with a dotted line have not been specifically examined. Conventional numbering of the carbon atoms is given for neurosporene and α -carotene. Inverted triangles (∇) mark positions of the double bonds introduced as a consequence of the desaturation reactions.

Figure 3 depicts the carotene endgroups which are found in plants.

Figure 4 is a DNA sequence and the predicted amino acid sequence of a lycopene ε-cyclase cDNA isolated from *A. thaliana* (SEQ ID NOS: 1 and 2). These sequences were deposited under Genbank accession number U50738. This cDNA is incorporated into the plasmid pATeps.

Figure 5 is a DNA sequence encoding the β -carotene hydroxylase isolated from A. thaliana (SEQ ID NO: 3). This cDNA is incorporated into the plasmid pATOHB.

Figure 6 is an alignment of the predicted amino acid sequences of A. thaliana β-carotene hydroxylase (SEQ ID NO: 4) with those of the bacterial β-carotene hydroxylase enzymes from Alicalgenes sp. (SEQ ID NO: 5) (Genbank D58422), Erwinia herbicola Eho10 (SEQ ID NO.: 6) (GenBank M872280), Erwinia uredovora (SEQ ID NO.: 7) (GenBank D90087) and Agrobacterium aurianticum (SEQ ID NO.: 8) (GenBank D58420). A consensus sequence is also shown. All five genes are identical where a capital letter appears in the consensus. A lowercase letter indicates that three of five, including A. thaliana, have the identical residue. TM; transmembrane.

Figure 7 is a DNA sequence of a cDNA encoding an IPP isomerase isolated from A. thaliana (SEQ ID NO: 9). This cDNA is incorporated into the plasmid pATDP5.

Figure 8 is a DNA sequence of a second cDNA encoding another IPP isomerase isolated from A. thaliana (SEQ ID NO: 10). This cDNA is incorporated into the plasmid pATDP7.

Figure 9 is a DNA sequence of a cDNA encoding an IPP isomerase isolated from *Haematococcus pluvialis* (SEQ ID NO: 11). This cDNA is incorporated into the plasmid pHP04.

Figure 10 is a DNA sequence of a second cDNA encoding another IPP isomerase isolated from *Haematococcus pluvialis* (SEQ ID NO: 12). This cDNA is incorporated into the plasmid pHP05.

Figure 11 is an alignment of the amino acid sequences predicted by IPP isomerase cDNAs isolated from A. thaliana (SEQ ID NO.: 16 and 18), H. pluvialis (SEQ ID NOS.: 14

5

10

15

20

25

30

10

15

20

25

30

and 15), Clarkia breweri (SEQ ID NO.: 17) (See, Blanc & Pichersky, Plant Physiol. (1995) 108:855; Genbank accession no. X82627) and Saccharomyces cerevisiae (SEQ ID NO.: 19) (Genbank accession no. J05090).

Figure 12 is a DNA sequence of the cDNA encoding an IPP isomerase isolated from *Tagetes erecta* (marigold; SEQ ID NO: 13). This cDNA is incorporated into the plasmid pPMDP1. xxx's denote a region not originally sequenced. Figure 21A shows the complete marigold sequence.

Figure 13 is an alignment of the consensus sequence of four plant β -cyclases (SEQ ID NO.: 20) with the A. thaliana lycopene ϵ -cyclase (SEQ ID NO.: 21). A capital letter in the plant β consensus is used where all four β -cyclase genes predict the same amino acid residue in this position. A small letter indicates that an identical residue was found in three of the four. Dashes indicate that the amino acid residue was not conserved and dots in the sequence denote a gap. A consensus for the aligned sequences is given, in capital letters below the alignment, where the β - and ϵ -cyclases have the same amino acid residue. Arrows indicate some of the conserved amino acids that will be used as junction sites for construction of chimeric cyclases with novel enzymatic activities. Several regions of interest including a sequence signature indicative of a dinucleotide-binding motif and two predicted transmembrane (TM) helical regions are indicated below the alignment and are underlined.

Figure 14 shows the nucleotide (SEQ ID NO:22) and amino acid sequences (SEQ ID NO:23) of the *Adonis palaestina* (pheasant's eye) ϵ -cyclase cDNA #5.

Figure 15A shows the nucleotide (SEQ ID NO:24) and amino acid sequences (SEQ ID NO:25) of a potato ε-cyclase cDNA. Figure 15B shows the amino acid sequence (SEQ ID NO:26) of a chimeric lettuce/potato lycopene ε-cyclase. Amino acids in lower case are from the lettuce cDNA and those in upper case are from the potato cDNA. The product of this chimeric cDNA has e-cyclase activity and converts lycopene to the monocyclic δ-carotene.

Figure 16 shows a comparison between the amino acid sequences of the *Arabidopsis* ϵ -cyclase (SEQ ID NO:27) and the potato ϵ -cyclase (SEQ ID NO:25).

Figure 17A shows the nucleotide sequence of the *Adonis palaestina* Ipi1 (SEQ ID NO:28) and Figure 17B shows the nucleotide sequence of the *Adonis palaestina* Ipi2 (SEQ ID NO: 29).

WO 99/63055 PCT/US99/12121

Figure 18A shows the nucleotide sequence of the *Haematoccus pluvialis* Ipi1 (SEQ ID NO:11) and Figure 18B shows the nucleotide sequence of the *Haematoccus pluvialis* Ipi2 (SEQ ID NO:30).

Figure 19A shows the nucleotide sequence of the Lactuca sativa (romaine lettuce)

Ipi1 (SEQ ID NO:31) and Figure 19B shows the nucleotide sequence of the Lactuca sativa

Ipi2 (SEQ ID NO: 32).

Figure 20 shows the nucleotide sequence of the *Chlamydomonas reinhardtii* Ipi1 (SEQ ID NO:33).

Figure 21A shows the nucleotide sequence of the *Tagetes erecta* (marigold) Ipi1 (SEQ ID NO:34) and Figure 21B shows the nucleotide sequence of the *Oryza sativa* (rice) Ipi1 (SEQ ID NO:35).

Figure 22 shows a amino acid sequence alignment of various plant and green algal isopentenyl isomerases (IPI) (SEQ ID NOS:16, 36-45).

Figure 23 shows a comparison between Adonis palaestina ϵ -cyclase cDNA #3 and cDNA #5 nucleotide sequences.

Figure 24 shows a comparison between *Adonis palaestina* ϵ -cyclase cDNA #3 and cDNA #5 predicted amino acid sequences.

Figure 25 shows a sequence alignment of various plant β - and ϵ -cyclases. Those sequences outlined in grey denote identical sequences among the ϵ -cyclases. Those sequences outlined in black denote identical sequences among both the β - and ϵ -cyclases.

Figure 26 shows a sequence alignment of the plant ϵ -cyclases from Figure 25. Those sequences outlined in black denote identical sequences among the ϵ -cyclases.

Figure 27 is a dendrogram or "tree" illustrating the degree of amino acid sequence similarity for various lycopene β - and ϵ -cyclases.

Figure 28 shows a comparison between Arabidopsis ϵ -cyclase and lettuce ϵ -cyclase predicted amino acid sequences.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention includes an isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity and having the amino acid sequence of SEQ ID NOS: 2, 4, 14-21, 23 or 25-27. Nucleic acids encoding lycopene ϵ -cyclase, β -carotene hydroxylase and IPP

5

10

15

20

25

30

PCT/US99/12121

ون مر

WO 99/63055

5

10

15

20

25

30

isomerases have been isolated from several genetically distant sources.

The present inventors have isolated nucleic acids encoding the enzyme IPP isomerase, which catalyzes the reversible conversion of isopentenyl pyrophosphate (IPP) to dimethylallyl pyrophosphate (DMAPP). IPP isomerase cDNAs were isolated from the plants A. thaliana, Tagetes erecta (marigold), Adonis palaestina (pheasant's eye), Lactuca sativa (romaine lettuce) and from the green algae H. pluvialis and Chlamydomonas reinhardtii. Alignments of the amino acid sequences predicted by some of these cDNAs are shown in Figures 12 and 22. Plasmids containing some of these cDNAs were deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville MD 20852 on March 4, 1996 under ATCC accession numbers 98000 (pHP05 - H. pluvialis); 98001 (pMDP1 - marigold); 98002 (pATDP7 - A. thaliana) and 98004 (pHP04 - H. pluvialis).

The present inventors have also isolated nucleic acids encoding the enzyme β-carotene hydroxylase, which is responsible for hydroxylating the β-endgroup in carotenoids. The nucleic acid of the present invention is shown in SEQ ID NO: 3 and Figure 5. The full length cDNA product hydroxylates both end groups of β-carotene as do products of cDNAs which encode proteins truncated by up to 50 amino acids from the N-terminus. Products of genes which encode proteins truncated between about 60-110 amino acids from the N-terminus preferentially hydroxylate only one ring. A plasmid containing this gene was deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville MD 20852 on March 4, 1996 under ATCC accession number 98003 (pATOHB - A. thaliana).

The present inventors have also isolated nucleic acids encoding the enzyme lycopene ϵ -cyclase, which is responsible for the formation of ϵ -endgroups in carotenoids. The A. thaliane ϵ -cyclase adds an ϵ ring to only one end of the symmetrical lycopene while the related β -cyclase adds a ring at both ends. The A. thaliana cDNA of the present invention is shown in Figure 4 and SEQ ID NO: 1. A plasmid containing this gene was deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville MD 20852 on March 4, 1996 under ATCC accession number 98005 (pATeps - A. thaliana).

In addition, lycopene ϵ -cyclases have been identified in lettuce and in *Adonis* palaestina (cDNA #5) which encode enzymes that convert lycopene to the bicyclic ϵ -carotene (ϵ , ϵ -carotene). An additional cDNA from *Adonis palaestina* (cDNA #3) encodes a lycopene ϵ -cyclase which converts lycopene into δ -carotene (ϵ , ψ -carotene) and differs from the lycopene ϵ -cyclase which forms bicyclic ϵ -carotene (ϵ , ϵ -carotene) by only 5 amino acids.

10

15

20

25

30

One or more of these amino acids may be modified by alteration of the nucleotide sequence in the #5 cDNA to obtain an enzyme which forms the bicyclic ϵ,ϵ -carotene. The sequences of the *Adonis palaestina* and *Arabidopsis thaliana* ϵ -cyclases have about 70% nucleotide identity and about 72% amino acid identity.

Initial experiments by the inventors with chimeric genes indicated that the part of the ϵ -cyclase which is responsible for adding 2ϵ rings to form ϵ, ϵ -carotene is the carboxy terminal portion of the gene. The lettuce ϵ -cyclase adds two ϵ rings to form ϵ, ϵ -carotene. A DNA encoding a partial potato ϵ -cyclase (missing its amino terminal portion), when combined with an amino terminal region from the lettuce ϵ -cyclase gene, produces a monocyclic δ -carotene (ϵ, ψ -carotene). With the discovery of the differences between the Adonis palaestina clone #3 and clone #5, the specific amino acids responsible for the addition of an extra ϵ ring have been identified (Figure 24). Specifically, amino acid 55 is Thr in clone #3 and Ser in clone #5, amino acid 210 is Asn in clone #3 and Asp in clone #5, amino acid 231 is Asp in clone #3 and Glu in clone #5, amino acid 352 is Ile in clone #3 and Val in clone #5, and amino acid 524 is Lys in clone #3 and Arg in clone #5. It can be appreciated that these changes are quite conservative, as only one change, at amino acid 210, changes the charge of the protein.

Thus, it is clear that the nucleic acids of the invention encoding the enzymes as presently disclosed may be altered to increase a particularly desirable property of the enzyme, to change a property of the enzyme, or to diminish an undesirable property of the enzyme. Such modifications can be by deletion, substitution, or insertion of one or more amino acids, and can be performed by routine enzymatic manipulation of the nucleic acid encoding the enzyme (such as by restriction enzyme digestion, removal of nucleotides by mung bean nuclease or *Bal31*, insertion of nucleotides by Klenow fragment, and by religation of the ends), by site-directed mutagenesis, or may be accidental, such as by low fidelity PCR or those obtained through mutations in hosts that are producers of the enzymes. These techniques as well as other suitable techniques are well known in the art.

Mutations can be made in the nucleic acids of the invention such that a particular codon is changed to a codon which codes for a different amino acid. Such a mutation is generally made by making the fewest nucleotide changes possible. A substitution mutation of this sort can be made to change an amino acid in the resulting protein in a non-conservative manner (i.e., by changing the codon from an amino acid belonging to a grouping

10

15

20

25

30

of amino acids having a particular size or characteristic to an amino acid belonging to another grouping) or in a conservative manner (i.e., by changing the codon from an amino acid belonging to a grouping of amino acids having a particular size or characteristic to an amino acid belonging to the same grouping). Such a conservative change generally leads to less change in the structure and function of the resulting protein. A non-conservative change is more likely to alter the structure, activity or function of the resulting protein. The present invention should be considered to include sequences containing conservative changes which do not significantly alter the activity or binding characteristics of the resulting protein.

The following is one example of various groupings of amino acids:

Amino acids with nonpolar R groups: Alanine, Valine, Leucine, Isoleucine, Proline, Phenylalanine, Tryptophan and Methionine.

Amino acids with uncharged polar R groups: Glycine, Serine, Threonine, Cysteine, Tyrosine, Asparagine and Glutamine.

Amino acids with charged polar R groups (negatively charged at Ph 6.0): Aspartic acid and Glutamic acid.

Basic amino acids (positively charged at pH 6.0): Lysine, Arginine and Histidine.

Another grouping may be those amino acids with phenyl groups: Phenylalanine, Tryptophan and Tyrosine.

Another grouping may be according to molecular weight (i.e., size of R groups). Particularly preferred substitutions are:

- Lys for Arg and vice versa such that a positive charge may be maintained;
- Glu for Asp and vice versa such that a negative charge may be maintained;
- Ser for Thr such that a free -OH can be maintained; and
- Gln for Asn such that a free NH₂ can be maintained.

Amino acid substitutions may also be introduced to substitute an amino acid with a particularly preferable property. For example, a Cys may be introduced to provide a potential site for disulfide bridges with another Cys. A His may be introduced as a particularly "catalytic" site (i.e., His can act as an acid or base and is the most common amino acid in biochemical catalysis). Pro may be introduced because of its particularly planar structure, which induces β -turns in the protein's structure.

It is clear that certain modifications of SEQ ID NOS: 2, 4, 14-21, 23 or 25-27 can take place without destroying the activity of the enzyme. It is noted especially that truncated

10

15

20

25

30

versions of the nucleic acids of the invention are functional. For example, several amino acids (from 1 to about 120) can be deleted from the N-terminus of the lycopene €-cyclases of the invention, and a functional protein can still be produced. This fact is made especially clear from Figure 25, which shows a sequence alignment of several plant €-cyclases. As can be seen from Figure 25, there is an enormous amount of sequence disparity between amino acid sequences 2 to about 50-70 (depending on the particular sequence, since gaps are present). There is less, but also a substantial amount of, sequence dissimilarity between about 50-70 to about 90-120 (depending on the particular sequence). Thereafter, the sequences are fairly conserved, except for small pockets of dissimilarity between about 275-295 to about 285-305 (depending on the particular sequence), and between about 395-415 to about 410-430 (depending on the particular sequence).

The present inventors have found that the amount of the 5' region present in the nucleic acids of the invention can alter the activity of the enzyme. Instead of diminishing activity, truncating the 5' region of the nucleic acids of the invention may result in an enzyme with a different specificity. Thus, the present invention relates to nucleic acids and enzymes encoded thereby which are truncated to within 0-50, preferably 0-25, codons of the 5' initiation codon of their prokaryotic counterparts as determined by alignment maps as discussed below.

For example, when the cDNA encoding A. thaliana β -carotene hydroxylase was truncated, the resulting enzyme catalyzed the formation of β -cryptoxanthin as the major product and zeaxanthin as minor product; in contrast to its normal production of zeaxanthin.

The present invention is intended to include those nucleic acid and amino acid sequences in which substitutions, deletions, additions or other modifications have taken place, as compared to SEQ ID NOS: 2, 4, 14-21, 23 or 25-27, without destroying the activity of the enzyme. Preferably, the substitutions, deletions, additions or other modifications take place at the 5' end, or any other of those positions which already show dissimilarity between any of the presently disclosed amino acid sequences (see also Figure 25) or other amino acid sequences which are known in the art and which encode the same enzyme (i.e., lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase).

In each case, nucleic acid and amino acid sequence similarity and identity is measured using sequence analysis software, for example, the Sequence Analysis, Gap, or BestFit software packages of the Genetics Computer Group (University of Wisconsin Biotechnology



Center, 1710 University Avenue, Madison, Wisconsin 53705), MEGAlign (DNAStar, Inc., 1228 S. Park St., Madison, Wisconsin 53715), or MacVector (Oxford Molecular Group, 2105 S. Bascom Avenue, Suite 200, Campbell, California 95008). Such software uses algorithms to match similar sequences by assigning degrees of identity to various substitutions, deletions, and other modifications, and includes detailed instructions as to useful parameters, etc., such that those of routine skill in the art can easily compare sequence similarities and identities. An example of a useful algorithm in this regard is the algorithm of Needleman and Wunsch, which is used in the Gap program discussed above. This program finds the alignment of two complete sequences that maximizes the number of matches and minimizes the number of gaps. Another useful algorithm is the algorithm of Smith and Waterman, which is used in the BestFit program discussed above. This program creates an optimal alignment of the best segment of similarity between two sequences. Optimal alignments are found by inserting gaps to maximize the number of matches using the local homology algorithm of Smith and Waterman.

15

10

5

Conservative (i.e. similar) substitutions typically include substitutions within the following groups: glycine and alanine; valine, isoleucine and leucine; aspartic acid, glutamic acid, asparagine and glutamine; serine and threonine; lysine and arginine; and phenylalanine and tyrosine. Substitutions may also be made on the basis of conserved hydrophobicity or hydrophilicity (see Kyte and Doolittle, *J. Mol. Biol.* 157: 105-132 (1982)), or on the basis of the ability to assume similar polypeptide secondary structure (see Chou and Fasman, *Adv. Enzymol.* 47: 45-148 (1978)).

25

20

If comparison is made between nucleotide sequences, preferably the length of comparison sequences is at least 50 nucleotides, more preferably at least 60 nucleotides, at least 75 nucleotides or at least 100 nucleotides. It is most preferred if comparison is made between the nucleic acid sequences encoding the enzyme coding regions necessary for enzyme activity. If comparison is made between amino acid sequences, preferably the length of comparison is at least 20 amino acids, more preferably at least 30 amino acids, at least 40 amino acids or at least 50 amino acids. It is most preferred if comparison is made between the amino acid sequences in the enzyme coding regions necessary for enzyme activity.

30

It should be appreciated that also within the scope of the present invention are nucleic acid sequences encoding lycopene ϵ -cyclases, IPP isomerases and β -carotene hydroxylases

10

15

20

25

30

which code for enzymes having the same amino acid sequence as SEQ ID NOS: 2, 4, 14-21, 23 or 25-27, but which are degenerate to the nucleic acids specifically disclosed herein.

The amino acid residues described herein are preferred to be in the "L" isomeric form. However, residues in the "D" isomeric form can be substituted for any L-amino acid residue, as long as the desired functional property of immunoglobulin-binding is retained by the polypeptide.

In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Sambrook et al, "Molecular Cloning: A Laboratory Manual" (1989); "Current Protocols in Molecular Biology" Volumes I-III [Ausubel, R. M., ed. (1994)]; "Cell Biology: A Laboratory Handbook" Volumes I-III [J. E. Celis, ed. (1994))]; "Current Protocols in Immunology" Volumes I-III [Coligan, J. E., ed. (1994)]; "Oligonucleotide Synthesis" (M.J. Gait ed. 1984); "Nucleic Acid Hybridization" [B.D. Hames & S.J. Higgins eds. (1985)]; "Transcription And Translation" [B.D. Hames & S.J. Higgins, eds. (1984)]; "Animal Cell Culture" [R.I. Freshney, ed. (1986)]; "Immobilized Cells And Enzymes" [IRL Press, (1986)]; B. Perbal, "A Practical Guide To Molecular Cloning" (1984).

The present invention also includes vectors. Suitable vectors according to the present invention comprise a nucleic acid of the invention encoding an enzyme involved in carotenoid biosynthesis or metabolism and a suitable promoter for the host, and can be constructed using techniques well known in the art (for example Sambrook et al., Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; Ausubel et al., Current Protocols in Molecular Biology, Greene Publishing and Wiley Interscience, New York, 1991). Suitable vectors for eukaryotic expression in plants are described in Frey et al., Plant J. (1995) 8(5):693 and Misawa et al, 1994a; incorporated herein by reference. Suitable vectors for prokaryotic expression include pACYC184, pUC119, and pBR322 (available from New England BioLabs, Bevery, MA) and pTrcHis (Invitrogen) and pET28 (Novagen) and derivatives thereof. The vectors of the present invention can additionally contain regulatory elements such as promoters, repressors, selectable markers such as antibiotic resistance genes, etc.

The nucleic acids encoding the carotenoid enzymes as described above, when cloned into a suitable expression vector, can be used to overexpress these enzymes in a plant

10

15

20

25

30

expression system or to inhibit the expression of these enzymes. For example, a vector containing the gene encoding lycopene ϵ -cyclase can be used to increase the amount of α -carotene and carotenoids derived from α -carotene (such as lutein and α -cryptoxanthin) in an organism and thereby alter the nutritional value, pharmacology and visual appearance value of the organism.

Therefore, the present invention includes a method of producing or enhancing the production of a carotenoid in a host cell, relative to an untransformed host cell, the method comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence to produce the protein.

The present invention also includes a method of modifying the production of carotenoids in a host cell, the method comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which produces an RNA and/or encodes for a protein which modifies lycopene ϵ -cyclase, IPP isomerase or β -carotene hydroxylase enzyme activity, relative to an untransformed host cell, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence in the host cell to modify the production of the carotenoids in the host cell, relative to the untransformed host cell.

The term "modifying the production" means that the amount of carotenoids produced in the host cell can be enhanced, reduced, or left the same, as compared to the untransformed host cell. In accordance with one embodiment of the present invention, the make-up of the carotenoids (i.e., the specific carotenoids produced) is changed vis a vis each other, and this change in make-up may result in either a net gain, net loss, or no net change in the total amount of carotenoids produced in the cell. In accordance with another embodiment of the present invention, the production or the biochemical activity of the carotenoids (or the enzymes which catalyze their formation) is enhanced by the insertion of an enzyme-encoding nucleic acid of the invention. In yet another embodiment of the invention, the production or the biochemical activity of the carotenoids (or the enzymes which catalyze their formation) may be reduced or inhibited by a number of different approaches available to those skilled in the art, including but not limited to such methodologies or approaches as anti-sense (e.g.,

WO 99/63055 PCT/US99/12121

Gray et al (1992) Plant Mol. Biol. 19:69-87), ribozymes (e.g., Wegener et al (1994) Mol. Gen. Genet. 245:465-470), co-suppression (e.g., Fray and Grierson (1993) Plant Mol. Biol. 22:589-602), targeted disruption of the gene (e.g., Schaefer et al. (1997) Plant J. 11:1195-1206), intracellular antibodies (e.g., Rondon and Marasco (1997) Ann. Rev. Microbiol. 51:257-283) or whatever other approaches rely on the knowledge or availability of the nucleic acid or amino acid sequences of the invention and/or portions thereof, to thereby reduce accumulation of carotenoids with ϵ rings and compounds derived from them (for ϵ -cyclase inhibition), or carotenoids with hydroxylated β rings and compounds derived from them (for β -hydroxylase inhibition), or, in the case if IPP isomerase, accumulation of any isoprenoid compound.

Preferably, at least a portion of the nucleic acid sequences used in the methods, vectors and host cells of the invention codes for an enzyme having an amino acid sequence which is at least 85% identical, preferably at least 90%, at least 95% or completely identical to SEQ ID NOS: 2, 4, 14-21, 23 or 25-27. Sequence identity is determined as noted above. Preferably, sequence additions, deletions or other modifications are made as indicated above, so as to not affect the function of the particular enzyme.

In a preferred embodiment, vectors are manufactured which contain a DNA encoding a eukaryotic IPP isomerase upstream of a DNA encoding a second eukaryotic carotenoid enzyme. The inventors have discovered that inclusion of an IPP isomerase gene increases the supply of substrate for the carotenoid pathway; thereby enhancing the production of carotenoid endproducts, as compared to a host cell which is not transformed with such a vector. This is apparent from the much deeper pigmentation in carotenoid-accumulating colonies of *E. coli* which also contain one of the aforementioned IPP isomerase genes when compared to colonies that lack this additional IPP isomerase gene. Similarly, a vector comprising an IPP isomerase gene can be used to enhance production of any secondary metabolite of dimethylallyl pyrophosphate and/or isopentenyl pyrophosphate (such as isoprenoids, steroids, carotenoids, etc.). The term "isoprenoid" is intended to mean any member of the class of naturally occurring compounds whose carbon skeletons are composed, in part or entirely, of isopentyl C₅ units. Preferably, the carbon skeleton is of an essential oil, a fragrance, a rubber, a carotenoid, or a therapeutic compound, such as paclitaxel.

A vector containing the cDNA encoding a lycopene ϵ -cyclase of the invention, preferably the lettuce lycopene ϵ -cyclase or Adonis ϵ -cyclase #5, can be used to increase the

5

10

15

20

25

30

10

15

20

25

30



amount of bicyclic €-carotene in an organism and thereby alter the nutritional value, pharmacology and visual appearance value of the organism. In addition, the transformed organism can be used in the formulation of therapeutic agents, for example in the treatment of cancer (see Mayne et al (1996) FASEB J. 10:690-701; Tsushima et al (1995) Biol. Pharm. Bull. 18:227-233).

An antisense strand of a nucleic acid of the invention can be inserted into a vector. For example, the lycopene ϵ -cyclase gene can be inserted into a vector and incorporated into the genomic DNA of a host, thereby inhibiting the synthesis of ϵ , β -carotenoids (lutein and α -carotene) and enhancing the synthesis of β , β -carotenoids (zeaxanthin and β -carotene).

The present invention also relates to novel enzymes which are encoded by the amino acid sequences of the invention, or portions thereof.

The present invention also relates to novel enzymes which can transform known carotenoids into novel or uncommon products. Currently ϵ -carotene (see Figure 2) and γ -carotene are commonly produced only in minor amounts. As described below, an enzyme can be produced which transforms lycopene to γ -carotene and lycopene to ϵ -carotene. With these products in hand, bulk synthesis of other carotenoids derived from them are possible. For example, ϵ -carotene can be hydroxylated to form lactucaxanthin, an isomer of lutein (one ϵ and one β ring) and zeaxanthin (two β rings) where both endgroups are, instead, ϵ rings.

In addition to novel enzymes produced by truncating the 5' region of known enzymes, as discussed above, novel enzymes which can participate in the formation of unusual carotenoids can be formed by replacing portions of one gene with an analogous sequence from a structurally related gene. For example, β -cyclase and ϵ -cyclase are structurally related (see Figure 13). By replacing a portion of β -lycopene cyclase with the analogous portion of ϵ -cyclase, an enzyme which produces γ -carotene will be produced (one β endgroup). Further, by replacing a portion of the lycopene ϵ -cyclase with the analogous portion of β -cyclase, an enzyme which produces ϵ -carotene will be produced (with some exceptions, such as the lettuce ϵ -cyclase, plant ϵ -cyclases normally produce a compound with one ϵ -endgroup, δ -carotene). Similarly, β -hydroxylase could be modified to produce enzymes of novel function by creation of hybrids with ϵ -hydroxylase.

Host systems according to the present invention can comprise any organism that already produces carotenoids or which has been genetically modified to produce carotenoids.

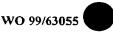
10

15

20

25

30



The IPP isomerase genes are more broadly applicable for enhancing production of any product dependent on DMAPP and/or IPP as a precursor.

Organisms which already produce carotenoids include plants, algae, some yeasts, fungi and cyanobacteria and other photosynthetic bacteria. Transformation of these hosts with vectors according to the present invention can be done using standard techniques such as those described in Misawa et al., (1990) supra; Hundle et al., (1993) supra; Hundle et al., (1991) supra; Misawa et al., (1991) supra; Sandmann et al., supra; and Schnurr et al., supra.

Transgenic organisms can be constructed which include the nucleic acid sequences of the present invention (Bird et al, 1991; Bramley et al, 1992; Misawa et al, 1994a; Misawa et al, 1994b; Cunningham et al, 1993). The incorporation of these sequences can allow the controlling of carotenoid biosynthesis, content, or composition in the host cell. These transgenic systems can be constructed to incorporate sequences which allow for the overexpression of the nucleic acids of the present invention. Transgenic systems can also be constructed containing antisense expression of the nucleic acid sequences of the present invention. Such antisense expression would result in the accumulation of the substrates of the substrates of the enzyme encoded by the sense strand.

A method for screening for eukaryotic genes which encode enzymes involved in carotenoid biosynthesis comprises transforming a prokaryotic host with a nucleic acid which may contain a eukaryotic or prokaryotic carotenoid biosynthetic gene; culturing said transformed host to obtain colonies; and screening for colonies exhibiting a different color than colonies of the untransformed host.

Suitable hosts include E. coli, cyanobacteria such as Synechococcus and Synechocystis, alga and plant cells. E. coli are preferred.

In a preferred embodiment, the above "color complementation" screening protocol can be enhanced by using mutants which are either (1) deficient in at least one carotenoid biosynthetic gene or (2) overexpress at least one carotenoid biosynthetic gene. In either case, such mutants will accumulate carotenoid precursors.

Prokaryotic and eukaryotic DNA or cDNA libraries can be screened in total for the presence of genes of carotenoid biosynthesis, metabolism and degradation. Preferred organisms to be screened include photosynthetic organisms.

10

15

20

25

30



E. coli can be transformed with these eukaryotic cDNA libraries using conventional methods such as those described in Sambrook et al, 1989 and according to protocols described by the vendors of the cloning vectors.

For example, the cDNA libraries in bacteriophage vectors such as lambdaZAP (Stratagene) or lambda ZIPLOX (Gibco BRL) can be excised en masse and used to transform *E.coli*.

Transformed E. coli can be cultured using conventional techniques. The culture broth preferably contains antibiotics to select and maintain plasmids. Suitable antibiotics include penicillin, ampicillin, chloramphenicol, etc. Culturing is typically conducted at 15-40°C, preferably at room temperature or slightly above (18-28°C), for 12 hours to 7 days.

Cultures are plated and the plates are screened visually for colonies with a different color than the colonies of the host $E.\ coli$ transformed with the empty plasmid cloning vector. For example, $E.\ coli$ transformed with the plasmid, pAC-BETA (described below), produce yellow colonies that accumulate β -carotene. After transformation with a cDNA library, colonies which contain a different hue than those formed by $E.\ coli/pAC$ -BETA would be expected to contain enzymes which modify the structure or accumulation of β -carotene. Similar $E.\ coli$ strains can be engineered which accumulate earlier products in carotenoid biosynthesis, such as lycopene, γ -carotene, etc.

Having generally described this invention, a further understanding can be obtained by reference to certain specific examples which are provided herein for purposes of illustration only and are not intended to be limiting unless otherwise specified.

EXAMPLE

I. <u>Isolation of β-carotene hydroxylase</u>

Plasmid Construction

An 8.6kb BgIII fragment containing the carotenoid biosynthetic genes of *Erwinia herbicola* was first cloned in the BamHI site of plasmid vector pACYC184 (chloramphenicol resistant), and then a 1.1kb BamHI fragment containing the *E. herbicola* β-carotene hydroxylase (*CrtZ*) was deleted. *E.coli* strains containing the resulting plasmid, pAC-BETA, accumulate β-carotene and form yellow colonies (Cunningham et al., 1994).

A full length cDNA encoding IPP isomerase of *Haematococcus pluvialis* (HP04) was first excised with *BamH*I and *Kpn*I from pBluescript SK-, and then ligated into the

WO 99/63055 PCT/US99/12121

corresponding sites of the pTrcHisA vector with high-level expression from the *trc* promoter (Invitrogen, Inc.). A fragment containing the IPP isomerase and *trc* promoter was subsequently excised with *EcoRV* and *KpnI*, treated with the Klenow fragment of DNA polymerase to produce blunt ends, and ligated in the Klenow-treated *HindIII* site of pAC-BETA. *E.coli* cells transformed with this new plasmid pAC-BETA-04 form orange colonies on LB plates (*vs.* yellow for those containing pAC-BETA) and cultures accumulate substantially more β-carotene (*ca.* two fold) than those that contain pAC-BETA.

Screening of an Arabidopsis cDNA Library

Several λ cDNA expression libraries of *Arabidopsis* were obtained from the *Arabidopsis* Biological Resource Center (Ohio State University, Columbus, OH) (Kieber et al., 1993). The λ cDNA libraries were excised *in vivo* using Stratagene's ExAssist SOLR system to produce a phagemid cDNA library wherein each phagemid contained also a gene conferring resistance to the antibiotic ampicillin.

E.coli strain DH10BZIP was chosen as the host cell for the screening and pigment production, although we have also used TOP10F' and XL1-Blue for this purpose. DH10B cells were transformed with plasmid pAC-BETA-04 and were plated on LB agar plates containing chloramphenicol at 50 µg/ml (from United States Biochemical Corporation). The phagemid Arabidopsis cDNA library was then introduced into DH10B cells already containing pAC-BETA-04. Transformed cells containing both pAC-BETA-04 and Arabidopsis cDNA library phagemids were selected on chloramphenicol plus ampicillin (150 µg/ml) agar plates. Maximum color development occurred after 3 to 7 days incubation at room temperature, and the rare bright yellow colonies were selected from a background of many thousands of orange colonies on each agar plate. Selected colonies were inoculated into 3 ml liquid LB medium containing ampicillin and chloramphenicol, and cultures were incubated at room temperature for 1-2 days, with shaking. Cells were then harvested by centrifugation and extracted with acetone in microfuge tubes. After centrifugation, the pigmented extract was spotted onto silica gel thin-layer chromatography (TLC) plates, and developed with a hexane:ether (1:1, by volume) mobile phases. B-carotene hydroxylaseencoding cDNAs were identified based on the appearance of a yellow pigment that comigrated with zeaxanthin on the TLC plates.

5

10

15

20

25

30

10

15

20

25

30



Subcloning and Sequencing

The plasmid containing the β-carotene hydroxylase cDNA was recovered and analyzed by standard procedures (Sambrook et al., 1989). The *Arabidopsis* β-carotene hydroxylase was sequenced completely on both strands on an automatic sequencer (Applied Biosystems, Model 373A, Version 2.0.1S). The cDNA insert of 0.95kb also was excised and ligated into the a pTrcHis vector. A *Bgl*II restriction site within the cDNA was used to remove that portion of the cDNA that encodes the predicted polypeptide N terminal sequence region that is not also found in bacterial β-carotene hydroxylases (Figure 6). A BglII-XhoI fragment was directionally cloned in BamHI-XhoI digested TrcHis vectors.

Pigment Analysis

A single colony was used to inoculate 50 ml of LB containing ampicillin and chloramphenicol in a 250-ml flask. Cultures were incubated at 28°C for 36 hours with gentle shaking, and then harvested at 5000 rpm in an SS-34 rotor. The cells were washed once with distilled H₂O and resuspended with 0.5 ml of water. The extraction procedures and HPLC were essentially as described previously (Cunningham et al, 1994).

II. Isolation and biochemical analysis of an Arabidopsis lycopene ϵ -cyclase Plasmid Construction

Construction of plasmids pAC-LYC, pAC-NEUR, and pAC-ZETA is described in Cunningham et al., (1994). In brief, the appropriate carotenoid biosynthetic genes from *Erwinia herbicola*, *Rhodobacter capsulatus*, and *Synechococcus* sp. strain PCC7942 were cloned in the plasmid vector pACYC184 (New England BioLabs, Beverly, MA). Cultures of *E. coli* containing the plasmids pAC-ZETA, pAC-NEUR, and pAC-LYC, accumulate ζ-carotene, neurosporene, and lycopene, respectively. The plasmid pAC-ZETA was constructed as follows: an 8.6-kb BgIII fragment containing the carotenoid biosynthetic genes of *E. herbicola* (GenBank M87280; Hundle et al., 1991) was obtained after partial digestion of plasmid pPL376 (Perry et al., 1986; Tuveson et al., 1986) and cloned in the BamHI site of pACYC184 to give the plasmid pAC-EHER. Deletion of adjacent 0.8- and 1.1-kb BamHI-BamHI fragments (deletion Z in Cunningham et al., 1994), and of a 1.1 kB SalI-SalI fragment (deletion X) served to remove most of the coding regions for the *E. herbicola* β-carotene hydroxylase (crtZ gene) and zeaxanthin glucosyltransferase (crtX gene), respectively. The

resulting plasmid, pAC-BETA, retains functional genes for geranylgeranyl pyrophosphate synthase (crtE), phytoene synthase (crtB), phytoene desaturase (crtI), and lycopene cyclase (crtY). Cells of *E. coli* containing this plasmid form yellow colonies and accumulate β -carotene. A plasmid containing both the lycopene ϵ - and β -cyclase cDNAs of *A. thaliana* was constructed by excising the ϵ -cyclase in clone y2 as a PvuI-PvuII fragment and ligating this piece in the SnaBI site of a plasmid (pSPORT 1 from GIBCO-BRL) that already contained the β -cyclase (Cunningham et al., 1996).

Organisms and Growth Conditions

E. coli strains TOP10 and TOP10 F' (obtained from Invitrogen Corporation, San Diego, CA) and XL1-Blue (Stratagene) were grown in Luria-Bertani (LB) medium (Sambrook et al., 1989) at 37°C in darkness on a platform shaker at 225 cycles per min. Media components were from Difco (yeast extract and tryptone) or Sigma (NaCl). Ampicillin at 150 μg/mL and/or chloramphenicol at 50 μg/mL (both from United States Biochemical Corporation) were used, as appropriate, for selection and maintenance of plasmids.

Mass Excision and Color Complementation Screening of an A. thaliana cDNA Library

A size-fractionated 1-2 kB cDNA library of A. thaliana in lambda ZAPII (Kieber et al., 1993) was obtained from the Arabidopsis Biological Resource Center at The Ohio State University (stock number CD4-14). Other size fractionated libraries were also obtained (stock numbers CD4-13, CD4-15, and CD4-16). An aliquot of each library was treated to cause a mass excision of the cDNAs and thereby produce a phagemid library according to the instructions provided by the supplier of the cloning vector (Stratagene; E. coli strain XL1-Blue and the helper phage R408 were used). The titre of the excised phagemid was determined and the library was introduced into a lycopene-accumulating strain of E. coli TOP10 F' (this strain contained the plasmid pAC-LYC) by incubation of the phagemid with the E. coli cells for 15 min at 37°C. Cells had been grown overnight at 30°C in LB medium supplemented with 2% (w/v) maltose and 10 mM MgSO₄ (final concentration), and harvested in 1.5 ml microfuge tubes at a setting of 3 on an Eppendorf microfuge (5415C) for 10 min. The pellets were resuspended in 10 mM MgSO₄ to a volume equal to one-half that of the

5

10

15

20

25



initial culture volume. Transformants were spread on large (150 mm diameter) LB agar petri plates containing antibiotics to provide for selection of cDNA clones (ampicillin) and maintenance of pAC-LYC (chloramphenicol). Approximately 10,000 colony forming units were spread on each plate. Petri plates were incubated at 37 C for 16 hr and then at room temperature for 2 to 7 days to allow maximum color development. Plates were screened visually with the aid of an illuminated 3x magnifier and a low power stage-dissecting microscope for the rare, pale pinkish-yellow to deep-yellow colonies that could be observed in the background of pink colonies. A colony color of yellow or pinkish-yellow was taken as presumptive evidence of a cyclization activity. These yellow colonies were collected with sterile toothpicks and used to inoculate 3ml of LB medium in culture tubes with overnight growth at 37°C and shaking at 225 cycles/min. Cultures were split into two aliquots in microfuge tubes and harvested by centrifugation at a setting of 5 in an Eppendorf 5415C microfuge. After discarding the liquid, one pellet was frozen for later purification of plasmid DNA. To the second pellet was added 1.5 ml EtOH, and the pellet was resuspended by vortex mixing, and extraction was allowed to proceed in the dark for 15-30 min with occasional remixing. Insoluble materials were pelleted by centrifugation at maximum speed for 10 min in a microfuge. Absorption spectra of the supernatant fluids were recorded from 350-550 nm with a Perkin Elmer lambda six spectrophotometer.

Analysis of isolated clones

20

5

10

15

Eight of the yellow colonies contained β -carotene indicating that a single gene product catalyzes both cyclizations required to form the two β endgroups of the symmetrical β -carotene from the symmetrical precursor lycopene. One of the yellow colonies contained a pigment with the spectrum characteristic of δ -carotene, a monocyclic carotenoid with a single ϵ endgroup. Unlike the β cyclase, this ϵ -cyclase appears unable to carry out a second cyclization at the other end of the molecule.

25

The observation that ϵ -cyclase is unable to form two cyclic ϵ -endgroups (e.g. the bicyclic ϵ -carotene) illuminates the mechanism by which plants can coordinate and control the flow of substrate into carotenoids derived from β -carotene versus those derived from α -carotene and also can prevent the formation of carotenoids with two ϵ endgroups.

30

The availability of the A. thaliana gene encoding the ϵ -cyclase enables the directed manipulation of plant and algal species for modification of carotenoid content and

composition. Through inactivation of the ϵ -cyclase, whether at the gene level by deletion of the gene or by insertional inactivation or by reduction of the amount of enzyme formed (by such as antisense technology), one may increase the formation of β -carotene and other pigments derived from it. Since vitamin A is derived only from carotenoids with β endgroups, an enhancement of the production of β -carotene versus α -carotene may enhance nutritional value of crop plants. Reduction of carotenoids with ϵ -endgroups may also be of value in modifying the color properties of crop plants and specific tissues of these plants. Alternatively, where production of α -carotene, or pigments such as lutein that are derived from α -carotene, is desirable, whether for the color properties, nutritional value or other reason, one may overexpress the ϵ -cyclase or express it in specific tissues. Wherever agronomic value of a crop is related to pigmentation provided by carotenoid pigments the directed manipulation of expression of the ϵ -cyclase gene and/or production of the enzyme may be of commercial value.

The predicted amino acid sequence of the A. thaliana ϵ -cyclase enzyme was determined. A comparison of the amino acid sequences of the β - and ϵ -cyclase enzymes of Arabidopsis thaliana (Fig. 13) as predicted by the DNA sequence of the respective cDNAs (Fig. 4 for the ϵ -cyclase cDNA sequence), indicates that these two enzymes have many regions of sequence similarity, but they are only about 37% identical overall at the amino acid level. The degree of sequence identity at the DNA base level, only about 50%, is sufficiently low such that we and others have been unable to detect this gene by hybridization using the β cyclase as a probe in DNA gel blot experiments.

REFERENCES

Each reference cited in this application and/or listed below is hereby incorporated by reference.

- 25 Bird et al, 1991 Biotechnology 9, 635-639.
 - Bishop et al., (1995) FEBS Lett. 367, 158-162.
 - Bramley, P.M. (1985) Adv. Lipid Res. 21, 243-279.
 - Bramley, P.M. (1992) Plant J. 2, 343-349.
 - Britton, G. (1988). Biosynthesis of carotenoids. In Plant Pigments, T.W. Goodwin, ed. (London: Academic Press), pp. 133-182.

30

5

10

15

20

15

25



Britton, G. (1979) Z. Naturforsch. Section C Biosci. 34, 979-985.

Britton, G. (1995) UV/Visible spectroscopy. In Carotenoids, Vol. IB: Spectroscopy,

G. Britton, S. Liaaen-Jensen, H.P. Pfander, eds. (Basel: Birkhauser Verlag), pp. 13-62.

Bouvier et al., (1994) Plant J. 6, 45-54.

Cunningham et al., (1985) Photochem. Photobiol. 42: 295-307.

Cunningham et al., (1993) FEBS Lett. 328, 130-138.

Cunningham et al., (1994) Plant Cell 6, 1107-1121.

Cunningham et al., (1996) Plant Cell 8, 1613-1626.

Davies, B.H. (1976). Carotenoids. In Chemistry and Biochemistry of Plant Pigments,

Vol. 2, T.W. Goodwin, ed (New York: Academic Press), pp. 38-165.

Del Sal et al., (1988). Nucl. Acids Res. 16, 9878.

Demmig-Adams & Adams, (1992) Ann. Rev. Plant Physiol. Mol. Biol. 43, 599-626.

Enzell & Back, (1995) Mass spectrometry. In Carotenoids, Vol. IB: Spectroscopy, G.

Britton, S. Liaaen-Jensen, H.P. Pfander, eds. (Basel: Birkhauser Verlag), pp. 261-320.

Frank & Cogdell (1993) Photochemistry and function of carotenoids in photosynthesis. In Carotenoids in Photosynthesis. A. Young and G. Britton, eds. (London: Chapman and Hall). pp. 253-326.

Goodwin, T.W. (1980). The Biochemistry of the Carotenoids. 2nd ed, Vol. 1 (London: Chapman and Hall.

20 Horvath et al., (1972) Phytochem. 11, 183-187.

Hugueney et al., (1995) Plant J. 8, 417-424.

Hundle et al., (1991) Photochem. Photobiol. 54, 89-93.

Jensen & Jensen, (1971) Methods Enzymol. 23, 586-602.

Kargl & Quackenbush, (1960) Archives Biochem. Biophys. 88, 59-63.

Kargl et al., (1960) Proc. Am. Hort. Soc. 75, 574-578.

Kieber et al., (1993) Cell 72, 427-441.

Koyama, Y. (1991) J. Photochem. Photobiol., B, 9, 265-80.

Krinsky, N.I. (1987) Medical uses of carotenoids. In Carotenoids, N.I. Krinsky, M.M. Mathews-Roth, and R.F. Taylor, eds. (New York: Plenum), pp. 195-206.

30 Kyte & Doolittle, (1982) J. Mol. Biol. 157, 105-132.

LaRossa & Schloss, (1984) J. Biol. Chem. 259, 8753-8757.

Misawa et al., (1994a) Plant J. 6, 481-489.

 $\#^{\Sigma}$

Misawa et al., (1994b) J. Biochem, Tokyo, 116, 980-985.

Norris et al., (1995) Plant Cell 7, 2139-2149.

Pecker et al., (1996) Submitted to Plant Mol. Biol.

Perry et al., (1986) J. Bacteriol. 168, 607-612.

Persson & Argos, (1994) J. Mol. Biol. 237, 182-192.

Plumley & Schmidt, (1987) Proc. Nat. Acad. Sci. USA 83, 146-150.

Plumley & Schmidt, (1995) Plant Cell 7, 689-704.

Rossmann et al., (1974) Nature 250, 194-199.

Rock & Zeevaart (1991) Proc. Nat. Acad. Sci. USA 88, 7496-7499.

Rost et al., (1995) Protein Science 4, 521-533.

Sambrook et al., (1989) Molecular Cloning: A Laboratory Manual, 2nd edition (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Sancar, A. (1994) Biochemistry 33, 2-9.

Sander & Schneider, (1991) Proteins 9, 56-68.

Sandmann, G. (1994) Eur. J. Biochem. 223, 7-24.

Scolnik & Bartley, (1995) Plant Physiol. 108, 1342.

Siefermann-Harms, D. (1987) Physiol. Plant. 69, 561-568.

Spurgeon & Porter, (1980). Biosynthesis of carotenoids. In Biochemistry of Isoprenoid Compounds, J.W. Porter, and S.L. Spurgeon, eds. (New York: Wiley), pp. 1-122.

Tomes, M.L. (1963) Bot. Gaz. 124, 180-185.

Tomes, M.L. (1967) Genetics 56, 227-232.

Tuveson et al., (1986) J. Bacteriol. 170, 4675-4680.

Van Beeumen et al., (1991) J. Biol. Chem. 266, 12921-12931.

Weedon & Moss, (1995) Structure and Nomenclature. In Carotenoids, Vol. IB:

Spectroscopy, G. Britton, S. Liaaen-Jensen, H.P. Pfander, eds. (Basel: Birkhauser Verlag), pp. 27-70.

Wierenga et al., (1986) J. Mol. Biol. 187, 101-107.

Zechmeister, L. (1962) Cis-Trans Isomeric Carotenoids, Vitamins A and Arylpolyenes. Springer-Verlag, Vienna.

Having now fully described the invention, it will be apparent to one of ordinary skill in the art that many changes and modifications can be made thereto without departing from the spirit or scope of the invention as set forth herein.

5

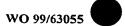
15

20

30

We claim:

- 1. An isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase enzyme activity and has an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23 or 25-27.
- 5 2. The nucleic acid sequence of claim 1, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23 or 25-27.
 - 3. A vector comprising the nucleic acid sequence of claim 1, wherein the nucleic acid sequence is operably linked to a promoter.
 - 4. A host cell which contains the vector of claim 3.
- 5. The host cell of claim 4, wherein the host cell is selected from the group consisting of a bacterial cell, an algal cell, a yeast cell and a plant cell.
 - 6. The host cell of claim 4, wherein the host cell is a photosynthetic cell.
- An isolated and/or purified protein having lycopene ε-cyclase enzyme activity and having an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23 or 25-27.
 - 8. The protein of claim 7, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23 or 25-27.



AMENDED CLAIMS

[received by the International Bureau on 15 November 1999 (15.11.99); original claims 1,2,7 and 8 amended; remaining claims unchanged (1 page)]

- 1. An isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ϵ -cyclase enzyme activity and has an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23, 25 or 26.
- 5 2. The nucleic acid sequence of claim 1, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23, 25 or 26.
 - 3. A vector comprising the nucleic acid sequence of claim 1, wherein the nucleic acid sequence is operably linked to a promoter.
 - 4. A host cell which contains the vector of claim 3.
- The host cell of claim 4, wherein the host cell is selected from the group consisting of a bacterial cell, an algal cell, a yeast cell and a plant cell.
 - 6. The host cell of claim 4, wherein the host cell is a photosynthetic cell.
- An isolated and/or purified protein having lycopene ε-cyclase enzyme activity and having an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23, 25 or 26.
 - 8. The protein of claim 7, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23, 25 or 26.

F I G. 1

3
$$3 \frac{2}{4} \frac{6}{5}$$

$$\beta \text{ cyclase}$$

$$3 \frac{6}{4} \frac{7}{5}$$

$$\beta \text{ endgroup}$$

$$\beta \text{ endgroup}$$

$$\epsilon \text{ endgroup}$$

±2 ±1

4/45

FIG.4A

FIG. 4B

FIG. 4

FIG. IIA

FIG.IIB

FIG. 11

FIG. 13A

FIG. 13B

FIG. 13

FIG.14A

FIG. 14B

FIG. 14

FIG. 22A

FIG. 22 B

FIG. 22

FIG. 4A

Arabidopsis thaliana epsilon cyclase:

| | | | | | aca | aaa | gga | aat | aat | tag | att | .cct | ctt | tct | gct | tgc | tat | acc | ttg | aca | 48 |
|-----|----------|-----------|-----------|-----------------|-----------------|-----------|----------|----------|-----------------|----------|-----------|-----------|-----------|----------|----------|----------|-----------------|-----------------|-----------|----------|-----|
| | gaa | caa | cat | aac | aat | ggt | gta | agt | ctt | ctç | gct | gta | ttc | gaa | att | att | tgg | agġ. | agga | aac | 108 |
| 1 | atg M | gag E | tgt: | gtt V | 999 G | gct A | agg R | aat N | ttc F | gca A | gca A | atg M | gcg A | gtt V | tca S | aca T | ttt F | ccg P | tcat S | tgg W | 168 |
| 21 | agt S | tgt C | cga R | agg R | aaa K | ttt F | cca P | gtg V | gct V | aag K | aga R | tac Y | agc S | tat Y | agg R | aat N | att I | cgc R | ttc F | ggt G | 228 |
| 41 | ttg L | tgt. C | agt S | gtc V | aga R | gct A | agc S | ggc G | ggc G | gga G | agt .S | tcc S | ggt G | agt S | gag K | agt S | tgt C | gta V | gcg A | gtg V | 288 |
| 61 | aga R | gaa S | gat D | ttc F | gct A | gac D | gaa E | gaa E | gat D | ttt F | gcg V | iaaa E | gct A | ggc G | ggt G | tct S | gag R | att I | cta L | ttt F | 348 |
| 81 | gtt V | caa Q | atg M | cag Q | cag Q | aac M | aaa K | gat D | atg M | gat D | gaa S | .cag Q | tct S | aag K | ctt L | gtt V | gat D | aag K | ttg(L | ct P | 408 |
| 101 | cct P | ata I | tca S | act I | ggt G | gat D | ggt G | gct A | ttg L | gat D | cat K | gtg V | gtt V | act I | ggc G | tgt C | ggt G | cct P | gcto A | ggt G | 468 |
| 121 | tta L | gcc A | ttg L | gct A | gca A | gaa K | tca S | gct A | aag K | ctt L | gga G | tta L | aaa K | gtt V | gga G | ctc L | att I | ggt. G | ccag P | gat D | 528 |
| 141 | ctt L | cct P | ttt F | act T | aac M | aat M | tac Y | ggt G | gtt V | tgg M | gaa K | gat D | gaa K | ttc F | aat N | gat D | ctt L | 9 9 9 | ctg(| caa G | 588 |
| 161 | aaa K | tgt C | att I | gag K | cat K | gtt V | tgg W | aga R | gag S | act T | att I | gcg V | cac Y | ctg L | gat D | gat D | gac D | aag K | ccta P | att I | 648 |
| 181 | acc T | att I | ggc G | cgt R | gct A | tat Y | gga G | aga R | gtt V | agt S | cga R | icgt R | ttg L | ctc L | cat X | gag E | gag E | ctt L | ttg: L | agg R | 708 |
| 201 | agg R | rtgt C | gtc V | gag K | rtca S | ggt G | gtc V | tcg S | tac Y | ctt L | ago S | tcg S | aaa K | gtt V | gac D | agc S | ata I | aca T | gaaq E | gct A | 768 |
| 221 | tgt S | gat D | .ggc G | ctt | aga X | ctt L | gtt V | gct A | tgt C | gac D | gac D | aat M | aac M | gtc V | att I | .ccc | tgc C | agg X | ctt | gcc A | 828 |
| 241 | act T | gtt V | gct A | tct S | .gga G | igca A | gct A | tcg S | gga G | aag K | cto L | ttg L | icaa Q | tac Y | gaa X | gtt V | ggt G | gga G | ccta P | aga R | 888 |
| | ata | tat | acc | ĸaa | act | :aca | tac | :aac | ata | ıdad | att | :aac | iaca | caa | aat | aat | cca | tat. | gata | сса | 948 |

FIG. 4B

| 201 | VC V Q I A I G V X V X N S P Y D P | |
|---------|---|------|
| 281 | gatcaaatggttttcatggattacagagat tatactaacgagaaagttcggagcttagaa D Q M V P M D Y R D Y T M X X V R S L X | 1008 |
| 301 | gctgagtatccaacgtttctgtacgccatg cctatgacaaagtcaagactcttcttcgag A K Y P T F L Y A M P M T K S R L F F K | 1068 |
| 321 | gagacatgtttggcctcaaaagatgtcatg ccctttgatttgctaaaaacgaagctcatg K T C L A S K D V M P F D L L K T K L M | 1128 |
| 341 | ttaagattagacacactcggaattcgaatt ctaaagacttacgaagaggagtggtcctat I P V G G S L P N T X Q K N L A F G A A | 1188 |
| 361 | atcccagttggtggttccttgccaaacacc gaacaaaagaatctcgcctttggtgctgcc I P V G G S L P M T X Q K N L A F G A A | 1248 |
| 381 | gctagcatggtacatcccgcaacaggctat tcagttgtgagatctttgtctgaagctcca A S M V M P A T G Y S V V R S L S X A P | 1308 |
| 401 | aaacatgcatcagtcatcgcagagatacta agagaagaga | 1368 |
| 421 | aatatttcaagacaagcttaggatacttta tggccaccagaaaggaaaagacagagagca M I S R Q A W D T L W P P E R X R Q R A | 1428 |
| 441 | ttctttctctttggtcttgcactcagagtt caattcgataccgaaggcattagaagcttc F F L F G L A L I V Q F D T X G I R S F | 1488 |
| 461 | ttccgtactttcttccgccttccaaaatgg atgtggcaagggtttctaggatcaacatta FRTPFRLPKWMWQGFLGSTL | 1548 |
| 481 | acatcaggagatctcgttctctttgcttta tacatgttcgtcatttcaccaaacaatttg T S G D L V L F A L Y M P V I S P M M L | 1608 |
| 501 | agaaaaaggtctcattaatcatctcatctct gatccaaccggagcaaccatgataaaaacc R K G L I N W L I S D P T G A T M I K T | 1668 |
| 521 | tatctcaaagtatgatttacttaccaactc ttaggtttgtgtatatatatgccgatttat Y L K V | 1728 |
| | | 1700 |
| | ctgaataatcgatcaaagaatggtatgtgg gttactaggaagttggaaacaaacacgtat | 1788 |
| | agaatctaaggagtgatcgaaatggagacg gaaacgaaaagaaaa | 1848 |
| | ccgtggctagtg | 1868 |

FIG. 5

gctctttctc ctcctcct accgatttcc gactccgcct cccgaaatcc 51 ttatccggat tctctccgtc tcttcgattt aaacgctttt ctgtctgtta cgtcgtcgaa gaacggagac agaattctcc gattgagaac gatgagagac 101 cggagagcac gagctccaca aacgctatag acgctgagta tctggcgttg 151 201 cgtttggcgg agaaattgga gaggaagaaa tcggagaggt ccacttatct 251 aatcgctgct atgttgtcga gctttggtat cacttctatg gctgttatgg 301 ctgtttacta cagattctct tggcaaatgg agggaggtga gatctcaatg 351 ttggaaatgt ttggtacatt tgctctctct gttggtgctg ctgttggtat 401 ggaattctgg gcaagatggg ctcatagagc tctgtggcac gcttctctat 451 ggaatatgca tgagtcacat cacaaaccaa gagaaggacc gtttgagcta 501 aacgatgttt ttgctatagt gaacgctggt ccagcgattg gtctcctctc ttatggattc ttcaataaag gactcgttcc tggtctctgc tttggcgccg 551 601 ggttaggcat aacggtgttt ggaatcgcct acatgtttgt ccacgatggt 651 ctcgtgcaca agcgtttccc tgtaggtccc atcgccgacg tcccttacct 701 ccgaaaggtc gccgccgctc accagctaca tcacacagac aagttcaatg 751 gtgtaccata tggactgttt cttggaccca aggaattgga agaagttgga 801 ggaaatgaag agttagataa ggagattagt cggagaatca aatcatacaa aaaggcctcg ggctccgggt cgagttcgag ttcttgactt taaacaagtt 851 901 ttaaatccca aattctttt ttgtcttctg tcattatgat catcttaaga 951 cggtct

+2

8745

FIG. 7

ccacgggtcc gcctccccgt ttttttccga tccgatctcc ggtgccgagg 1 actcagctgt ttgttcgcgc tttctcagcc gtcaccatga ccgattctaa 51 cgatgctgga atggatgctg ttcagagacg actcatgttt gaagacgaat 101 151 gcattctcgt tgatgaaaat aatcgtgtgg tgggacatga cactaagtat 201 aactgtcatc tgatggaaaa gattgaagct gagaatttac ttcacagagc tttcagtgtg tttttattca actccaagta tgagttgctt ctccagcaac 251 ggtcaaaaac aaaggttact ttcccacttg tgtggacaaa cacttgttgc 301 agccatecte tttacegtga atecgagett attgaagaga atgtgettgg tgtaagaaat gccgcacaaa ggaagctttt cgatgagctc ggtattgtag 401 451 cagaagatgt accagtegat gagttcacte cettgggaeg catgetttae 501 aaggcacctt ctgatgggaa atggggagag cacgaagttg actatctact 551 cttcatcgtg cgggatgtga agcttcaacc aaacccagat gaagtggctg 601 agatcaagta cgtgagcagg gaagagctta aggagctggt gaagaaagca 651 gatgetggeg atgaagetgt gaaactatet ceatggttea gattggtggt 701 ggataatttc ttgatgaagt ggtgggatca tgttgagaaa ggaactatca 751 ctgaagctgc agacatgaaa accattcaca agctctgaac tttccataag 801 ttttggatct tccccttccc ataataaaat taagagatga gacttttatt 851 gattacagac aaaactggca acaaaatcta ttcctaggat ttttttttgc 901 tttttattta cttttgattc atctctagtt tagttttcat cttaaaaaaa 951 aaaa

FIG. 8

1 caccaatgte tgtttcttct ttatttaate teccattgat tegeeteaga 51 tototogoto titogtotto titttottot titocGATTTG CCCATCGTCC TCTGTCATCG ATTTCACCGA GAAAGTTACC GAATTTTCGT GCTTTCTCTG 101 GTACCGCTAT GACAGATACT AAAGATGCTG GTATGGATGC TGTTCAGAGA 151 201 CGTCTCATGT TTGAGGATGA ATGCATTCTT GTTGATGAAA CTGATCGTGT 251 TGTGGGGCAT GTCAGCAAGT ATAATTGTCA TCTGATGGAA AATATTGAAG 301 CCAAGAATTT GCTGCACAGG GCTTTTAGTG TATTTTTATT CAACTCGAAG 351 TATGAGTTGC TTCTCCAGCA AAGGTCAAAC ACAAAGGTTA CGTTCCCTCT AGTGTGGACT AACACTTGTT GCAGCCATCC TCTTTACCGT GAATCAGAGC 401 451 TTATCCAGGA CAATGCACTA GGTGTGAGGA ATGCTGCACA AAGAAAGCTT 501 CTCGATGAGC TTGGTATTGT AGCTGAAGAT GTACCAGTCG ATGAGTTCAC 551 TCCCTTGGGA CGTATGCTGT ACAAGGCTCC TTCTGATGGC AAATGGGGAG 601 AGCATGAACT TGATTACTTG CTCTTCATCG TGCGAGACGT GAAGGTTCAA CCAAACCCAG ATGAAGTAGC TGAGATCAAG TATGTGAGCC GGGAAGAGCT 651 701 GAAGGAGCTG GTGAAGAAAG CAGATGCAGG TGAGGAAGGT TTGAAACTGT 751 CACCATGGTT CAGATTGGTG GTGGACAATT TCTTGATGAA GTGGTGGGAT CATGTTGAGA AAGGAACTTT GGTTGAAGCT ATAGACATGA AAACCATCCA 801 851 CAAACTCTGA ACATCTTTTT TTAAAGTTTT TAAATCAATC AACTTTCTCT TCATCATTTT TATCTTTTCG ATGATAATAA TTTGGGATAT GTGAGACACT 901 TACAAAACTT CCAAGCACCT CAGGCAATAA TAAAGTTTGC GGCCGC

FIG. 9

| 1 | CTCGGTAGCT | GGCCACAATC | GCTATTTGGA | ACCTGGCCCG | GCGGCAGTCC |
|------|------------|------------|------------|------------|------------|
| 51 | GATGCCGCGA | TGCTTCGTTC | GTTGCTCAGA | GGCCTCACGC | ATATCCCCC |
| 101 | CGTGAACTCC | GCCCAGCAGC | CCAGCTGTGC | ACACGCGCGA | CTCCAGTTTA |
| 151 | AGCTCAGGAG | CATGCAGATG | ACGCTCATGC | AGCCCAGCAT | CTCAGCCAAT |
| 201 | CTGTCGCGCG | CCGAGGACCG | CACAGACCAC | ATGAGGGGTG | CAAGCACCTG |
| 251 | GGCXGGCGGG | CAGTCGCAGG | ATGAGCTGAT | GCTGAAGGAC | GAGTGCATCT |
| 301 | TGGTGGATGT | TGAGGACAAC | ATCACAGGCC | ATGCCAGCAA | GCTGGXGTGT |
| 351 | CACAAGTTCC | TACCACATCA | GCCTGCAGGC | CTGCTGCACC | GGGCCTTCTC |
| 401 | TGTGTTCCTG | TTTGACGATC | AGGGGCGACT | GCTGCTGCAA | CAGCGTGCAC |
| 451 | GCTCAAAAAT | CACCTTCCCA | AGTGTGTGGA | CGAACACCTG | CTGCAGCCAC |
| 501 | CCTTTACATG | GGCXGXCCCC | AGATGAGGTG | GACCAACTAA | GCCAGGTGGC |
| 551 | CGACGGAACA | GTACCTGGCG | CAAAGGCTGC | TGCCATCCGC | AAGTTGGAGC |
| 601 | ACGAGCTGGG | GATACCAGCG | CACCAGCTGC | CGGCAAGCGC | GTTTCGCTTC |
| 651 | CTCACGCGTT | TGCACTACTG | TGCCGCGGAC | GTGCAGCCAG | CTGCGACACA |
| 701 | ATCAGCGCTC | TGGGGCGAGC | ACGAAATGGA | CTACATCTTG | TTCATCCGGG |
| 751 | CCAACGTCAC | CTTGGCGCCC | AACCCTGACG | AGGTGGACGA | AGTCAGGTAG |
| 801 | GTGACGCAAG | AGGAGCTGCG | GCAGATGATG | CAGCCGGACA | ACGGGCTGCA |
| 851 | ATGGTCGCCG | TGGTTTCGCA | TCATCGCCGC | GCGCTTCCTT | GAGCGTTGGT |
| 901 | GGGCTGACCT | GGACGCGGCC | CTAAACACTG | ACAAACACGA | GGATTGGGGA |
| 951 | ACGGTGCATC | ACATCAACGA | AGCGTGÄAAG | CAGAAGCTGC | AGGATGTGAA |
| 1001 | GACACGTCAT | GGGGTGGAAT | TGCGTACTTG | GCAGCTTCGT | ATCTCCTTTT |
| 1051 | TCTGAGACTG | AACCTGCAGT | CAGGTCCCAC | AAGGTCAGGT | AAAATGGCTC |
| 1101 | GATAAAATGT | ACCGTCACTT | TTTGTCGCGT | ATACTGAACT | CCAAGAGGT |
| 1151 | ********* | **** | | | |

FIG. 10

| 1 | CICGGTAGCT | GGCCACAATC | GCTATTTGGA | ACCTGGCCCG | GCGGCAGTCC |
|------|------------|------------|------------|------------|------------|
| 51 | GATGCCGCGA | TGCTTCGTTC | GTTGCTCAGA | GGCCTCACGC | ATATCCCGCG |
| 101 | CGTGAACTCC | GCCCAGCAGC | CCAGCTGTGC | ACACGCGCGA | CTCCAGTTTA |
| 151 | AGCTCAGGAG | CATGCAGCTG | CTTTCCGAGG | ACCGCACAGA | CCACATGAGG |
| 201 | GGTGCAAGCA | CCTGGGCAGG | CGGGCAGTCG | CAGGATGAGC | TGATGCTGAA |
| 251 | GGACGAGTGC | ATCTTGGTAG | ATGTTGAGGA | CAACATCACA | GGCCATGCCA |
| 301 | GCAAGCTGGA | GTGTCACAAG | TTCCTACCAC | ATCAGCCTGC | AGGCCTGCTG |
| 351 | CACCGGGCCT | TCTCTGTGTT | CCTGTTTGAC | GATCAGGGGC | GACTGCTGCT |
| 401 | GCAACAGCGT | GCACGCTCAA | AAATCACCTT | CCCAAGTGTG | TGGACGAACA |
| 451 | CCTGCTGCAG | CCACCCTTTA | CATGGGCAGA | CCCCAGATGA | GGTGGACCAA |
| 501 | CTAAGCCAGG | TGGCCGACGG | AACAGTACCT | GGCGCAAAGG | CTGCTGCCAT |
| 551 | CCGCAAGTTG | GAGCACGAGC | TGGGGATACC | AGCGCACCAG | CTGCCGGCAA |
| 601 | GCGCGTTTCG | CTTCCTCACG | CGTTTGCACT | ACTGTGCCGC | GGACGTGCAG |
| 651 | CCAGCTGCGA | CACAATCAGC | GCTCTGGGGC | GAGCACGAAA | TGGACTACAT |
| 701 | CTTGTTCATC | CGGGCCAACG | TCACCTTGGC | GCCCAACCCT | GACGAGGTGG |
| 751· | ACGAAGTCAG | GTACGTGACG | CAAGAGGAGC | TGCGGCAGAT | GATGCAGCCG |
| 801 | GACAACGGGC | TTCAATGGTC | GCCGTGGTTT | CGCATCATCG | CCGCGCGCTT |
| 851 | CCTTGAGCGT | TGGTGGGCTG | ACCTGGACGC | GGCCCTAAAC | ACTGACAAAC |
| 901 | ACGAGGATTG | GGGAACGGTG | CATCACATCA | ACGAAGCGTG | AAGGCAGAAG |
| 951 | CTGCAGGATG | TGAAGACACG | TCATGGGGTG | GAATTGCGTA | CTTGGCAGCT |
| 1001 | TCGTATCTCC | TTTTTCTGAG | ACTGAACCTG | CAGAGCTAGA | GTCAATGGTG |
| 1051 | CATCATATTC | ATCGTCTCTC | TTTTGTTTTA | GACTAATCTG | TAGCTAGAGT |
| 1101 | CACTGATGAA | TCCTTTACAA | CTTTCAAAAA | | |

FIG. IIA

| HPO4 HPO5 ATDP7 C.brew. ATOP5 S.cerev. | MLRSLLRGLT MSVSSLFNLP MS.SSMLNFT | HIPRVNSAQQ HIPRVNSAQQ .LIRLRSLA. .ASRIVSLPL .TGPPPRFFP PHGAVSSYAK | PSCAHARLQF | AHRPLŠSIS. PLCFFSPISL | PRKLPNFRAF TQRFSAKLTF |
|---|--|--|--|--|--|
| | SEDRTDHMRG SGTA.MTD SSQATT.MGE T.MTD | | DELMLKDECI RRLMFEDECI | LVDVEDNITG LVDETDRVVG LVDENDKVVG LVDENNRVVG | HASKLECHKF HVSKYNCHLM HESKYNCHLM |
| | LPHOPAGLLH ENIEAKNLLH ENIESENLLH EKIEAENLLH | RAFSVFLFDD RAFSVFLFNS RAFSVFLFNS | QGRLLLQQRA KYELLLQQRS KYELLLQQRS KYELLLQQRS | NTKVTFPLVW ATKVTFPLVW | TNTCCSHPLH TNTCCSHPLY |
| | RE RE | SELIQDNALG SELIDENCLG SELIEENVLG | VRNAAQRKLL VRNAAQRKLL VRNAAQRKLF | HELGIPAHOL HELGIPAHOL DELGIVAEDV DELGIPAEDL DELGIVAEDV HELGIPEDET | PV.DEFTPLG PV.DEFTPLG PV.DEFTPLG |
| | 201 RLHYCAADVQ RLHYCAADVQ RMLY RILY RMLY | .KAPSDGKWG .KAPSDGKWG .KAPSDGKWG | EHEMDYILFI EHEMDYILFI EHELDYLLFI EHELDYLLFI EHEVDYLLFI EHEVDYLLFI EHEIDYILFY | RANVTLRANVTLVRDVKVIRDVNLVRDVKL KINAKENLTV | APNPDEVDEV QPNPDEVAEI DPNPDEVAEV QPNPDEVAEI |

÷ :

14/45

FIG. IIB

300
RYVTQEELRQ MMQ...PDN GLQWSPWFRI IAARFLERWW ADLDAALNTD
RYVTQEELRQ MMQ...PDN GLQWSPWFRI IAARFLERWW ADLDAALNTD
KYVSREELKE LVKKADAGEE GLKLSPWFRL VVDNFLMKWW DHVEKGTLVE
KYMNRDDLKE LLRKADAEEE GVKLSPWFRL VVDNFLFKWW DHVEKGSLKD
KYVSREELKE LVKKADAGDE AVKLSPWFRL VVDNFLMKWW DHVEKGTITE
KWVSPNDLKT MF....ADP SYKFTPWFKI ICENYLFNWW EQLDDLSEVE

301
KHEDWGTVHH INEA*
KHEDWGTVHH INEA*
A.IDMKTIHK L*
A.ADMKTIHK L*
A.ADMKTIHK L*
A.ADMKTIHK L*
NDRQ...IHR ML*

F I G. 12

| 1 | ccaaaaacaa | ctcaaatctc | ctccgtcgct | cttactccgc | catgggtgac |
|-----|------------|------------|---|------------|------------|
| 51 | | | tcagcgacgt | | |
| 101 | | | acaatgtggt | | |
| 151 | | | attgaaacag | | |
| 201 | | | ttcaaaatac | | |
| 251 | | | ttcctttagt | | |
| 301 | | | tccgagcttg | | |
| 351 | | | ××××××××××××××××××××××××××××××××××××××× | | |
| 401 | | | ××××××××××××××××××××××××××××××××××××××× | | |
| 451 | | | xxxxxxxxx | | |
| 501 | | | xxxxxxxxx | | |
| 551 | | | xxxxxxxxx | | |
| 601 | | | | | |
| 651 | | | xxxxxxxxxx | | |
| 701 | | | xxxxxxxxx | | |
| 751 | | | tgaaaaccat | | |
| 801 | | | agcctaataa | | |
| | | • | ttaacaactt | | |
| 851 | | | gtaagatttt | | |
| 901 | | ttgatggttt | gcaatttcaa | gttcctatcg | acatgtagtg |
| 951 | atctaaaaaa | | | | |

| vkS-f-s- kfGK- csgvc LCSVRASGGG SSGSESCVAV REDFADEXDF | DLAVVGGGPA GLAVAQQVSE AGLSVCSIDP DHVVIGCGPA GLALAAESAK LGLKVGLIGP DV-G-GPA GLA-AGL-VI-P | Dinucleotide-binding signature | 210 -t-KDL-RPY GRVNRKQLKS KMMOKCI-NG DKPITIGRAY GRVSRRLLHE ELLRRCVESG |
|---|---|-------------------------------------|--|
| 1 MECVGARNFA AMAVSTFPSW SCRRKFPVVK RYSYRNÍRFG LCSVRASGGG SSGSESCVAV REDFADEXDF | Cyanobacterial enzyme begins———————————————————————————————————— | Possible subunit interaction domain | 141 -PKLIWPNN YGVWVDEFEA MDLLDCLDaT WSGa-VYiDd -t-KDL-RPY GRVNRKQLKS KMTOKCI-NG DLPFTNN YGVWEDEFND LGLQKCIEHV WRETIVYLDD DKPITIGRAY GRVSRRLLHE ELLRRCVESG PNN YGVW-DEFLC WVY-DDR-Y GRV-RL |
| Plant <i>beta</i> A.t.epsilon Consensus | Plant <i>beta</i> A.t.epsilon Consensus | ã. | Plant <i>beta</i> A.t.epsilon Consensus |

Conserved region #1

F1G. 13 A

Predicted TM helix 13B Conserved region #3 Conserved region #4 Conserved region #5 Predicted TM helix Conserved region #2 品配合 Plant beta A.t.epsilon Consensus Plant beta A.t.epsilon Plant beta 4. t.epsilon Plant beta A.t.epsilon Consensus Consensus Consensus

FIG. 14A

Adonis palaestina ε-cyclase cDNA #5 Length: 1898

| 1 | addinantat totattaata ttactataaa attat |
|------|--|
| 51 | |
| 101 | . Gadacecede ecocolocie (CECTICARA ACARCARACE SAFAFARACE |
| 151 | - 3 3 |
| 201 | coolytyty dellelydda caagaaacet tagtagttea aaactagti |
| 251 | acadeataca tegatatudi ierrettata aantagattt teaactaaaa |
| 301 | 9009009909 GARACHUMAN LANAANTTOT ATTACTTATA 3303000+++ |
| | |
| 351 | cocadatyta ytaadtaadu (Claffinana aacaggeess getsssssi |
| 401 | addition caller in a second contract the con |
| 451 | agging agg conduction for an aggregation |
| 501 | - 99°°9°°°°°°°°°°°°°°°°°°°°°°°°°°°°°°°° |
| 551 | 99'9'9'9'9'9 adddCddll Caaadatett ddaettdaac gttgtatar |
| 601 | ged by the general additional additional actions and actions and actions and actions are actions as the second actions and actions are actions as the second actions are actions as the second actions are actions as the second actions are actions as the second actions are actions as the second actions are actions as the second actions are actions as the second actions are actions as the second actions are actions as the second actions are actions as the second actions are actions as the second actions are actions as the second actions are actions as the second actions as the second actions are actions as the second actions are actions as the second actions are actions as the second actions are actions as the second actions are actions as the second actions as the second actions are actions as a second action actions are actions as a second action actions are actions as a second action actions are actions as a second action actions are actions as a second action actions are actions as a second action actions are actions as a second action actions are actions as a second action actions are actions as a second action actions are actions as a second action actions are actions as a se |
| 651 | contributed by the contributed and the contrib |
| 701 | trace and the second of the se |
| 751 | gradagyate delyddycid gidaiggera fageettata atttataaaa |
| 801 | digagatett tateettige aggettaeta etattaeate tagagagaet |
| 851 | tedyyyada tittuydadta Idaadtadd doccotcoto +++o+o+o+o |
| 901 | dategettat yyddiddadd flaaddinna naacaateca tacgateeca |
| 951 | deridatyyi diildiddac Tacadadact atatocaaca gaaattacaa |
| 1001 | rycicyyddy ddyddidioc aacatttoti tatotcatoc coatgeogae |
| 1051 | adeadyacti tittilidadd adacetafff aaceteaaa gataceataa |
| 1101 | carregater arrivaddada adactratut carrattuaa gaetetagat |
| 1151 | diccadgita tadaddilla laaadaanaa tootcatata ttoototta |
| 1201 | - 19911CILLA CCADACACA ACCAAAAAAA CCTAGCA+++ GG+GC+GGA |
| 1251 | - caused user year ended design at a contract and at a contract and a contract an |
| 1301 | gaageteeda aatatuette tutaattuea aagatttaa agaaagaaa |
| 1351 | The contract of the contract o |
| 1401 | dayearyyay cayrerriga eeaaaaaaaaaaaaaaaaaaaaaaaaaa |
| 1451 | |
| 1501 | ducaticiti duddccttct tccacttace aacttaceta tactaceeti |
| 1551 | tooligyque iledelatea terrenate tentettatt ttacatata |
| 1601 | Arguilli LUUCUCCAAA CAGCATGAGG Atgtcactta taaaaaattt |
| 1651 | Secretary of the secretary and the secretary s |
| 1701 | agreedated accalidade letagraffi caccasatas atazagatas |
| 1751 | - conducting canalidated telegraphic tateetatae tetaataa |
| 1801 | Laddyladdi UCCUUUTTTA ATATTATTAT atcaaacca coost i |
| 1851 | aaagtaaatt tattgataca aaagtagttt ttttccttaa aaaaaaaa |
| | Jangua Coucocoda adadddd |

FIG. 14B

Adonis palaestina E-cyclase #5 predicted polypeptide TRANSLATE from: 113 to: 1702 Length: 529 amino acids

MELLGVRNLI SSCPVWTFGT RNLSSSKLAY NIHRYGSSCR VDFQVRADGG SGSRSSVAYK EGFVDEEDFI KAGGSELLFV QMQQTKSMEK QAKLADKLPP 51 101 IPFGESVMDL VVIGCGPAGL SLAAEAAKLG LKVGLIGPDL PFTNNYGVWE DEFKDLGLER CIEHAWKDTI VYLDNDAPVL IGRAYGRVSR HLLHEELLKR 151 201 CVESGVSYLD SKVERITEAG DGHSLVVCEN EIFIPCRLAT VASGAASGKL LEYEVGGPRV CVQTAYGVEV EVENNPYDPN LMVFMDYRDY MQQKLQCSEE 251 301 EYPTFLYVMP MSPTRLFFEE TCLASKDAMP FDLLKRKLMS RLKTLGIOVT 351 KVYEEEWSYI PVGGSLPNTE QKNLAFGAAA SMVHPATGYS VVRSLSEAPK YASVIAKILK QDNSAYVVSG QSSAVNISMQ AWSSLWPKER KRQRAFFLFG 401 LELIVOLDIE ATRTÉFRTFF RLPTWMWGF LGSSLSSFDL VLFSMYMFVL 451 501 APNSMRMSLV RHLLSDPSGA VMVRAYLER*

+ 1

20/45

FIG. 15A

DNA sequence of potato cDNA (GenBank R27545) obtained from Nicholas J. Provart

```
potato.seq Length: 1378 August 2, 1996 13:06 Type: N Check: 605
     tagcggnnnn naggatgagt tcaaagatct tggtcttcaa gcctgcattg
     aacatgtttg gcgggatacc attgtatatc ttgatgatga tgatcctatt
  51
     cttattggcc gtgcctatgg aagagttagt cgccatttac tgcacgagga
 101
 151
     gttactcaaa aggtgtgtgg aggcaggtgt tttgtatcta aactcgaaag
     tggataggat tgttgaggcc acaaatggcc acagtcttgt agagtgcgag
 201
     ggtgatgttg tgattccctg caggtttgtg actgttgcat cgggagcagc
 251
 301
     ctcggggaaa ttcttgcagt atgagttggg aggtcctaga gtttctgttc
 351
     aaacagctta tggagtggaa gttgaggtcg ataacaatcc atttgacccg
 401
     agcctgatgg ttttcatgga ttatagagac tatgtcagac acgacgctca
     atctttagaa gctaaatatc caacatttct ctatgccatg cccatgtctc
 451
 501
     caacacgagt ctttttcgag gaaacttgtt tggcttcaaa agatgcaatg
 551
     ccattcgatc tgttaaagaa aaaattgatg ttacgattga acaccctcgg
     tgtaagaatt aaagaaattt atgaggagga atggtcttac ataccagttg
601
651<sup>-</sup>
     gaggatettt gecaaataea gaacaaaaaa eacttgeatt tggtgetget
 701
     gctagcatgg ttcatccagc cacaggttat tcagtcgtca gatcactgtc
 751
     tgaagctcca aaatgcgcct tcgtgcttgc aaatatatta cgacaaaatc
     atagcaagaa tatgcttact agttcaagta ccccgagtat ttcaactcaa
801
     gcttggaaca ctctttggcc acaagaacga aaacgacaaa gatcgttttt
851
901
     cctatttgga ctggctctga tattgcagct ggatattgag gggataaggt
     catttttccg cgcgttcttc cgtgtgccaa aatggatgtg gcagggattt
 951
     cttggttcaa gtctttcttn agcagacctc atgttatttg ccttctacat
1001
     gtttattatt gcaccaaatg acatgagaag aggcttaatc agacatcttt
1051
1101
     tatctgatcc tactggtgca acattgataa gaacttatct tacattttag
1151
     agtaaattcc tcctacaata gttgttgaan nagaggcctc attacttcag
1201
     attcataaca gaaatcgcgg tctctcgagg ccttgtatat aacattttca
1251
     ctaggttaat attgcttgaa taagttgcac agtttcagtt tttgtatctg
1301
     cttcttttt gtccaagatc atgtattgan ccaatttata tacattgcca
1351
     gtatatataa attttataaa aaaaaaaa
```

poteps.pep Length: 378 TRANSLATE from: 14 to: 1147

- 1 DEFKDLGLQA CIEHVWRDTI VYLDDDDPIL IGRAYGRVSR HLLHEELLKR 51 CVEAGVLYLN SKVDRIVEAT NGHSLVECEG DVVIPCRFVT VASGAASGKF 101 LQYELGGPRV SVQTAYGVEV EVDNNPFDPS LMVFMDYRDY VRHDAQSLEA
- 151 KYPTFLYAMP MSPTRVFFEE TCLASKDAMP FDLLKKKLML RLNTLGVRIK
- 201 EIYEEEWSYI PYGGSLPNTE QKTLAFGAAA SMVHPATGYS VVRSLSEAPK 251 CAFVLANILR QNHSKNMLTS SSTPSISTQA WNTLWPQERK RORSFFLFGL
- 301 ALILQLDIEG İRSFFRAFFR VPKWMWQGFL GSSLSXADLM LFAFYMFIIA
- 351 PNDMRRGLIR HLLSDPTGAT LIRTYLTF*

FIG. 15B

Chimeric lettuce/potato lycopene $\epsilon\text{-cyclase}$: converts lycopene to $\delta\text{-}$ carotene, the lettuce cDNA converts lycopene to $\epsilon\text{-carotene}$ and the potato cDNA does not produce an active enzyme

(amino acids in lower case are from lettuce and those in uppercase are from the potato cDNA; an $Ava\Pi$ site in common to the two cDNAs was used to construct the chimera)

| 1 | mecfgarnmt | atmavftcpt | ftdcnirhkf | sllkqrrftn | lsassslrqi |
|-----|------------|-------------------|------------|------------|------------|
| 51 | kcsaksdrcv | vdkqgisvac | eedyvkaggs | elffvqmqrt | ksmesqskls |
| 101 | eklagipign | cildlvvigc | gpaglalaae | saklglnvgl | igpdlpftnn |
| 151 | ygvwqdefig | lglegciehs | wkdtlvyldd | adpirigray | grvhrdllhe |
| 201 | ellrrcvesg | vsylsskver | iteapngysl | iecegnitip | crlatvasga |
| 251 | asgkfleyel | gGPRVSVQTA | YGVEVEVDNN | PFDPSLMVFM | DYRDYVRHDA |
| 301 | QSLEAKYPTF | LYAMPMSPTR | VFFEETCLAS | KDAMPFDLLK | KKLMLRLNTL |
| 351 | GVRIKEIYEE | EWSYIPVGGS | LPNTEQKTLA | FGAAASMVHP | ATGYSVVRSL |
| 401 | SEAPKCAFVL | ANILRONHSK | NMLTSSSTPS | ISTQAWNTLW | PQERKRQRSF |
| 451 | FLFGLALILQ | LDIEGIRSFF | RAFFRVPKWM | WQGFLGSSLS | XADLMLFAFY |
| 501 | METTAPNOMR | RGI TRHLLSD | PTGATLIRTY | LTF* | |

FIG. 16

| blosi | mparison of Arabidopsis E-cyclase x potato E-cyclase (parti um62.cmp Gap Weight: 12 Average Match: 2, Length Weight: 4 Average Mismatch: -2, Quality: 1485 Length: | .912 .003 |
|----------|---|--------------|
| Match | Ratio: 3.929 Gaps: Percent Similarity: 79.893 Percent Identity: 76. display thresholds for the alignment(s): = IDENTITY := 2 . = 1 | 1 .139 |
| 151 1 | EDEFNDLGLQKCIEHVWRETIVYLDDDKPITIGRAYGRVSRRLLHEELLR | |
| 201 | RCVESGVSYLSSKVDSITEASDGLRLVACDDNNVIPCRLATVASGAASGK | 250 |
| 50 | . . : . | 99 |
| 251 | LLQYEVGGPRVCVQTAYGVEVEVENSPYDPDQMVFMDYRDYTNEKVRSLE | 300 |
| 100 | . | 149 |
| 301 | AEYPTFLYAMPMTKSRLFFEETCLASKDVMPFDLLKTKLMLRLDTLGIRI | 350 |
| 150 | | 199 |
| 351 | LKTYEEEWSYIPVGGSLPNTEOKNLAFGAAASMVHPATGYSVVRSLSEAP | 400 |
| 200 | KETYEEEWSYTPVGGSLPNTEQKTLAFGAAASMVHPATGYSVVRSLSEAP | 249 |
| 401 | | 445 |
| 250 | : : . | 299 |
| 446 | LALIVOFDTEGIRSFFRTFFRLPKWMWQGFLGSTLTSGDLVLFALYMFVI | 495 |
| 300 | | 349 |
| 496 | | |
| 350 | . .: : : : : APNDMRRGLIRHLLSDPTGATLIRTYLTF 378 | |

FIG. 17A

Adonis palaestina Ipil attcatcttc agcagcgctg tcgtactctt tctatatctt cttccatcac 51 taacagtagt cgccgacggt tgaatcggct attcgcctca acgtcaacta 101 tgggtgaagt cactgatgct ggaatggatg ctgttcagaa gcggctcatg ttcgacgacg aatgtatttt ggtggatgag aatgacaagg tcgtcgggca 151 201 tgattccaaa tacaactgtc atttgatgga aaagatagag gcagaaaatt 251 tgcttcacag agccttcagt gttttcttgt tcaactcaaa atatgaattg 301 cttcttcagc aacgatccgc cacaaaggta acattcccgc tcgtatggac 351 aaacacatgt tgcagtcatc ctctctttcg tgattccgag ctcatagaag 401 aaaattatct cggtgtacga aacgctgcac aaagaaagct tttagacgag 451 ctaggcattc cagctgaaga tgtcccagtt gatgaattta ctcctcttgg 501 tcqcattctt tacaaagctc catctgacgg caaatgggga gagcacgaat 551 tggactatct cctatttatt gtccgagatg tgaaatacga tccaaaccca 601 gatgaagttg ctgatgctaa gtatgttaat cgcgaggagt tgagagagat 651 actgagaaaa gctgatgctg gtgaagaggg actcaagttg tctccttggt ttagattggt tgttgataac tttttgttca agtggtggga tcatgtagag 701 751 cagggtacga ttaaggaagt tgctgacatg aaaactatcc acaagttgac ttaagaggac ttctctcctc tgttctacta tttgtttttt gctacaataa 801 gtgggtggtg ataagcagtt tttctgtttt ctttaattta tggcttttga 851 atttgcctcg atgttgaact tgtaacatat ttagacaaat atgagacctt 901 951 gtaagttgaa tttgaggctg aatttatatt tttgggaaca taataatgtt 1001 aa

- 1

24 / 45

FIG. 17B

Adonis palaestina Ipi2 ttttaaagct ctttcgctcc accaccatca aagccagcca aatttctctg tacaaaagtt aaaaacaccg ctttgggctt tggcccctcc atatcggaat 51 ccttgtttac gatacgcatc taaaccagta attctcggtt ttaatttgtt 101 tcctaaatta ggcccctttc cggaatcccg agaattatgt cgtcgatcag 151 gattaatcct ttatatagta tcttctccac caccactaaa acattatcag 201 cttcqtqttc ttctcccgct gttcatcttc agcagcgttg tcgtactctt 251 tctatttctt cttccatcac taacagtcct cgccgagggt tgaatcggct 301 gttcgcctca acgtcgacta tgggtgaagt cgctgatgct ggtatggatg 351 ccgtccagaa gcggcttatg ttcgacgatg aatgtatttt ggtggatgag 401 aatgacaagg tcgtcggaca tgattccaaa tacaactgtc atttgatgga 451 aaagatagag gcagaaaact tgcttcacag agccttcagt gttttcttat 501 tcaactcaaa atacgagttg cttcttcagc aacgatctgc aacgaaggta 551 acattecege tegtatggae aaacacetgt tgeageeate ecetetteeg 601 tgattccgaa ctcatagaag aaaattttct cggggtacga aacgctgcac 651 aaaggaagct tttagacgag ctaggcattc cagctgaaga cgtaccagtt 701 gatgaattca ctcctcttgg tcgcattctt tacaaagctc catctgacgg 751 aaaatgggga gagcacgaac tggactatct tctgtttatt gtccgagatg 801 tgaaatacga tccaaaccca gatgaagttg ctgacgctaa gtacgttaat 851 cgcgaggagt tgaaagagat actgagaaaa gctgatgcag gtgaagaggg 901 aataaagttg tctccttggt ttagattggt tgtggataac tttttgttca 951 agtggtggga tcatgtagag gaggggaaga ttaaggacgt cgccgacatg 1001 aaaactatcc acaagttgac ttaagagaaa gtctcttaag ttctactatt 1051 tggtttttgc ttcaataagt ggatggtgat gagcagtttt tatgcttcct 1101 ttaattttgg cttttcaatt tgctttatgt gttgaacttg taacatattt 1151 agtcaaatat gagaccttgt gagttgaatt tgaggttata tttatagttt 1201 tgggaacata aaaaaaaaa 1251

FIG. 18A

Haematococcus pluvialis Ipil 1 ctcggtagct ggccacaatc gctatttgga acctggcccg gcggcagtcc gatgccgcga tgcttcgttc gttgctcaga ggcctcacgc atatcccccg 51 cgtgaactcc gcccagcagc ccagctgtgc acacgcgcga ctccagttta 101 agctcaggag catgcagatg acgctcatgc agcccagcat ctcagccaat 151 ctgtcgcgcg ccgaggaccg cacagaccac atgaggggtg caagcacctg 201 ggcaggcggg cagtcgcagg atgagctgat gctgaaggac gagtgcatct 251 tggtggatgt tgaggacaac atcacaggcc atgccagcaa gctggagtgt 301 cacaagttcc taccacatca gcctgcaggc ctgctgcacc gggccttctc 351 tgtgttcctg tttgacgatc aggggcgact gctgctgcaa cagcgtgcac 401 gctcaaaaat caccttccca agtgtgtgga cgaacacctg ctgcagccac 451 cctttacatg ggcagacccc agatgaggtg gaccaactaa gccaggtggc 501 cgacggaaca gtacctggcg caaaggctgc tgccatccgc aagttggagc 551 acgagetggg gataccageg caccagetge eggeaagege gtttegette 601 ctcacgcgtt tgcactactg tgccgcggac gtgcagccag ctgcgacaca 651 atcagcgctc tggggcgagc acgaaatgga ctacatcttg ttcatccggg 701 ccaacgtcac cttggcgccc aaccctgacg aggtggacga agtcaggtac 751 gtgacgcaag aggagctgcg gcagatgatg cagccggaca acgggctgca 801 atggtcgccg tggtttcgca tcatcgccgc gcgcttcctt gagcgttggt 851 gggctgacct ggacgcggcc ctaaacactg acaaacacga ggattgggga 901 acggtgcatc acatcaacga agcgtgaaag cagaagctgc aggatgtgaa 951 gacacgtcat ggggtggaat tgcgtacttg gcagcttcgt atctcctttt 1001 tctgagactg aacctgcagt caggtcccac aaggtcaggt aaaatggctc 1051 gataaaatgt accgtcactt tttgtcgcgt atactgaact ccaagaggtc 1101 aaaaaaaaa aaaaa 1151

- 1

26/45

FIG. 18B

Haematococcus pluvialis Ipi2

```
tggaacctgg cccggcggca gtccgatgcc gcgatgcttc gttcgttgct
 51
     cagaggeete acgeatatee egegegtgaa eteegeecag cageecaget
101
     gtgcacacgc gcgactccag tttaagctca ggagcatgca gctgcttgcc
151
     gaggaccgca cagaccacat gaggggtgca agcacctggg caggcgggca
201
     gtcgcaggat gagctgatgc tgaaggacga gtgcatctta gtggatgctg
251
     acgacaacat cacaggccat gccagcaagc tggagtgcca caaattccta
301
     ccacatcage etgeaggeet getgeacegg geettetetg tgtteetgtt
351
     tgacgaccag gggcgactgc tgctgcaaca gcgtgcacgc tcaaaaatca
401
     ccttcccaag tgtgtggacg aacacctgct gcagccaccc tctacatggg
451
     cagaccccag atgaggtgga ccaactaagc caggtggccg acggcacagt
501
     acctggcgca aaagctgctg ccatccgcaa gttggagcac gagctgggga
551
     taccagcgca ccagctgccg gcaagcgcgt ttcgcttcct cacgcgtttg
601
     cactactgtg ccgcggacgt gcagccggct gcgacacaat cagcgctctg
651
      gggcgagcac gagatggact acatcttatt catccgggcc aacgtcacct
701
     tggcgcccaa ccctgacgag gtggacgaag tcaggtacgt gacgcaagag
751
      gagetgegge agatgatgea geeggacaae gggttgeaat ggtegeegtg
     gtttcgcatc atcgccgcgc gcttccttga gcgttggtgg gctgacctgg
801
851
      acgcggccct aaacactgac aaacacgagg attggggaac ggtgcatcac
     atcaacgaag cgtgaaggca gaagctgcag gatgtgaaga cacgtcatgg
901
 951
      ggtggaattg cgtacttggc agcttcgtat ctcctttttc tgagactgaa
1001
      cctgcagage tagagtcaat ggtgcatcat attcatcgte tetetttgt
1051
      tttagactaa tctgtagcta gagtcactga tgaatccttt acaactttca
1101
      aaaaaaaa
```

•



Lactuca sativa Ipil tgccaaaatg ttgaaatttc ccccttttaa aaccattgct accatgatct cttctccata ttcttccttc ttgctgcctc ggaaatcttc tttccctcca atgccgtctc tcgcagccgc tagtgttttc ctccaccctc tttcgtctgc 101 cgčtałgggc gałtccagca tggałgctgt ccagcgacgt ctcałgttcg 151 atgacgaatg cattitiggtg gatgagaatg acaaagtggt tggccatgat actaaataca attgtcattt gatggagaag attgaaaagg gaaatatgct 201 251 acacagagca ttcagtgtgt tcttgttcaa ctcgaaatat gaattactcc 301 ttcagcaacg ttctgcaacc aaggtgactt tccctttggt atggacaaac 351 acqtqttqca qccatccact atacagggag agtgagctta ttgacgaaaa 401 cgcccttggg gtgaggaatg ctgcacagag gaagctcctg gatgaactcg gcatcctgg agcagatgtt ccggttgatg agttcactcc attgggtcgc 451 501 551 attctataca aggccgcatc ggatggaaag tggggagaac atgaacttga ttacctgctg tftatggtac gfgafgttgg tffggafccg aacccagafg 601 aagtgaaaga tgtaaaatat gtgaaccggg aagagctgaa ggaattggta aggaaggcgg atgctggtga agagggtgtg aagctgtccc cgtggttcaa attgattgtc gataatttct tgtttcagtg gtgggatcga ctccataagg 651 701 751 801 gaaccctaac cgaagctatt gatatgaaaa caatccacaa actcacataa ăaacactaca ctagtaggag ăgaggăttat atgagatatt tgttatatgt 851 gaaattgaaa ttcăgatgaă tgcttgtatt tatttctatt tggacaaact 901 tcaacttctt tttgctacct tatcagaaaa aaaaa 951

FIG. 19B

Lactuca sativa Ipi2 tattcgcttc aaaatctctt ccattaactg ctcaaatctc caccttcgcc ggtctťaatc tccgccggcg cactttcacc accataaccg ccgccatggg tgacgattcc ggcatggacg ctgtccagag acgtctcatg tttgatgatg 101 aătgčatttt ğğttgătgaă aatgacaatg ttcttgggca tgataccaaa 151 tacăattgtc ăcttgatgga gaagattgag aaagatăătt tgcttcatag agcattcagt gtattttat tcaattcaaa atacgaatta ctccttcagc 201 251 aaaggtcaga aaccaaggtg acatttcctt tggtatggac aaacacctgt 30ī tgcăğccatc cactatăcağ agaatcggag ttaattcccg aaaatgccct tggggtcaga aatgctgcac agaggaagct tctagatgaa ctcggtatcc 351 401 ctgctgaaga tgttccagtt gatgagttca caactttagg tcgcatgttg 451 tačaaggctc catctgatgg aaaatggggt gaacatgaag ttgattacct actcttcctc gtgcgtgacg ttgccgtgaa cccaaaccct gatgaggtgg 501 551 cggacattag atacgtgaac caagaagagt taaaagagtt actaaggaag 601 651 ğgtggacaac ttcttgttca aatggtggga tcatgtccaa aaggggacac 701 tcaatgaagc aattgacatg aaaaccattc ataagttgat atgaaaaatg 751 gttaatattt atggtggtgg tttggagcta ataatttgtg tgttcaagtc tcggtccttc ttttttaac gtttttttt tttcttttat tgggagtgtt 801 851 tattgtgtac ttgtaacgta ggccctttgg ttacgcttta agagtttaat 901 aaagaaccac cgttaattta aaaaaaaaa aaaaaaaa

FIG. 20

Chlamydomonas reinhardtii Ipil

(Note: the isomerase cDNA probably ends at ca. base 1103; the second half of the cDNA is similar to extensin and other hydroxyproline-rich structural proteins)

FIG. 21A

Tagetes erecta Ipil ccaaaaacaa ctcaaatctc ctccgtcgct cttactccgc catgggtgac gactccggca tggatgctgt tcagcgacgt ctcatgtttg acgatgaatg 101 cattttggtg gatgagtgtg acaatgtggt gggacatgat accaaataca 151 attgtcactt gatggagaag attgaaacag gtaaaatgct gcacagagca ttcagcgttt ttctattcaa ttcaaaatac gagttacttc ttcagcaacg 201 251 gtctgcaacc aaggtgacat ttcctttagt atggaccaac acctgttgca ğccatccact ctăcağagaa tccgagcttg ttcccgaaaa cgcccttgga gtaagaaatg ctgcacagag gaagctgttg gatgaactcg gtatccctgc 301 351 tgaagatgtt cccgttgatc agtttactcc tttaggtcgc atgctctaca 401 aggetecate tgatggaaag tggggagaac atgaaettga etacetaett tteatagtga gagaegttge tgtaaaeceg aaceeagatg aagtggegga 451 501 551 tatcaaatat gtganccang aagagttaaa ggagctgcta aggaaagcag atgcggggga ggagggtttg aagctgtctc catggttcag gttagtggtt 601 gataactict tgttcaagtg gtgggatcat gtgcaaaagg gtacactcac 651 tgaagcaatt gatatgaaaa ccatacacaa gctgatatag aaacacaccc 701 tčaačcgaaa agttcaagcc taataattcg ggttgggtcg ggtctaccat caattgttt tttctttaa gaagttttaa tctctatttg agcatgttga 751 801 851 ttcttgtctt ttgtgtgtaa gattttgggt ttcgtttcag ttgtaataat gaaccattga tggtttgcaa tttcaagttc ctatcgacat gtagtgatct 901 951 aaaaaa

FIG. 21B

Oryza sative Ipil cctccctttg cctcgcgcag aggcggccgc gccttctccg ccgcgaggat ggccggcgcc gccgccgcg tggaggacgc cgggatggac gaggtccaga agcggctcat gttcgacgac gaatgcattt tggtggatga acaagacaat 101 gttgttggcc atgaatcaaa atataactgc catctgatgg aaaaaatcga 201 atctgaaaat ctacttcata gggctttcag tgtattcctg ttcaactcaa aatafgaact cctactccag caacgatctg caacaaaggt tacatttcct ctagtttgga ccaacacttg ctgcagccat cctctgtacc gtgagtctga 251 301 351 gcttatacag gaaaactacc ttggtgttag aaatgctgct cagaggaagc tcttggatga gctgggcatc ccagctgaag atgtgccagt tgaccaattc acccctcttg gtcggatgct ttacaaggcc ccatctgatg gaaaatgggg 401 451 501 tgaacacgag čttgactacc tgctgttcat cgtccgcgac gtgaaggtag tčccgaačcč ggačgaagtg gčcgatgtga aatacgtgag čcgtgagcag ctgaaggagc tcatccgcaa agcggacgcc ggagaggaag gcctgaagct 551 601 gtčtcččtgg ttccggčtgg tfgffgačaa čftčcfčatg ggctggtggg 651 atcacgtcga gaaaggcacc ctcaacgagg ccgtggacat ggagaccatc cacaagctga agtaaggact gcgatgttgt ggctggaaag aatgatcctg 701 751 aagactctgt tcttgtgctg ctgcatatta ctcttaccag ggaagttgca 801 gaagtcagaa gaagcttttg tatgtttctg ggtttggagc ttggaagtgt tgggctctgc tgactgagag attcccttat agagtgtcta tgttaattta 851 901 gčaaacttčt atattataca tgattagtta attgttcggt gtctgaataa 951 āgaacaatag catgttccat gtttatttgc t 1001

| | ന | ء | ന | | 0 | V | _ | 0 | 7 | 6 | 60 | | | ١, | | | | | | | | | | |
|---------|---|---|---|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---|---|--|---|---------|---------------------------------|----------------------------------|----------------------------------|------------------------------------|------------------------------------|----------------------------|----------------------------|----------------------------|-----------------------------------|-----------------------------------|--------------------------------------|
| | 88 5 | 530 | 88 | 52 | 19 | 19 | 241 | <u>5</u> | 24 | 529 | 259 | | | _ | | | (e) | (e) | | | | | | |
| 270 | RKADA | KAUA A | RKADA | RKADA | RKADA | RKADA | KKADA | KKADA | d) | db | A | | | | ttuce, | ttuce, | t's e | t's e | | | | | | |
| | LKELL | LKELV | LKELL | LKEIL | LREIL | LKELI | LKELV | LKELV | LROW | LROM | LOSM | | • | (plo | ne le | ne le | easau | easan | | | | is | is | dtii |
| 556 | VSHEE | VNREE | VNQEE | VNREE | VNREE | VSREQ | VSREE | VSREE | VTQEE | VTQEE | VTLPE | | | Tagetes erecta (marigold) | Lactuca sativa (romaine lettuce) | Lactuca sativa (romaine lettuce, | Adonis palaestina (pheasant's eye) | Adonis palaestina (pheasant's eye) | ce) | iana | iana | Haematococcus pluvialis | Haematococcus pluvialis | Chlamydomonas reinhardtii |
| 255 256 | DIKY | OVKY | NDIRY | DAKY | IDAKY | IDVKY | LEIKY | \E1KY |)EVRY | EVRY |)ATRY | | | scta (| iva (| iva (| restir | restir | Oryza sativa (rice) | Arabidopsis thaliana | Arabidopsis thaliana | ld sno | ld sno | nas re |
| | 4PDEV# | PDEV | PDEV/ | IPDEV | IPDEV | IPDEV | PDEV | (PDEV | IPDEV | IPDEV | NPDEVI | | ٠. | es ere | sa sal | sa sal | s palc | s palc | sativ | sisdop | sisdop | tococ | tococ | лдошол |
| 241 | KAPSDGKWG EHELDYLLFIVRD VAVNPNPDEVADIKY VSHEELKELLRKADA | KAASDGKWG EHELDYLLFMVRD VGLDPNPDEVKDVKY VNREELKELVRKADA | KAPSDGKWG EHEVDYLLFLVRD VAVNPNPDEVADIRY VNQEELKELLRKADA | VKYDPNPDEVADAKY VNREELKEILRKADA | VKYDPNPDEVADAKY VNREELREILRKADA | VKVVPNPDEVADVKY VSREQLKELIRKADA | VKVQPNPDEVAEIKY VSREELKELVKKADA | VKLQPNPDEVAEIKY VSREELKELVKKADA | VTLAPNPDEVDEVRY VTQEELRQMMQP | VTLAPNPDEVDEVRY VTQEELROMMQP- | VPASSFSFLTRLHYC AADTATHG-PAAEWG EHEVDYVLFVRPQQP VSLQPNPDEVDATRY VTLPELQSMMA- | | | Taget | Lactu | Lactua | Adoni: | Adoni | Oryza | Arabi | Arabia | Наета | Наета | Chlam |
| 240 241 | Q | 8 9 | RD | 8- | Q | 8- | | | | | 900 | | | 232 | 280 | 529 | 562 | 34 | 238 | 58 4 | 233 | .93 | 202 | 307 |
| | LLFIV | LLFMV | LLFLV | LLFIV | LLFIV | LLFIV | LLFIV | LLFIV | /ILFIR | /ILFIR | NLFVR | | | | | | | | - | | | | | |
| 526 | EHELDY | EHELD | EHEVDY | KAPSDGKWG EHELDYLLFIVRD | KAPSDGKWG EHELDYLLFIVRD | KAPSDGKWG EHELDYLLFIVRD | KAPSDGKWG EHELDYLLFIVRD | KAPSDGKWG EHEVDYLLFIVRD | LPASAFRFLTRLHYC AADVOPAATQSALWG EHEMDYILFIRAN | I PASAFREI TRI HYC. AADVOPAATOSALWG EHEMDYILFIRAN | EHEVD | | 316 | DNFLFKMADHVQK GTLTEAIDMKTI HKLI | HKLT | HH | HKLT | 黑LT | HKLK | 形 | ₩ | ARFLERWADLDA ALNTDKHEDWGTV HHINEA | ARFLERWADLDA ALNTDKHEDWGTV HHINEA | TQPAFLPAWWGDLKR RWRPGGSRLSDWGTI HRVM |
| 225 226 | -KK | -KMC | -KMG | -886 | -KMG | -KFG | -KWG | -KMG | ALWG | ALWG | AEMG | | 315 316 | MKT1 | | | | | | | | WGTV | ₩CTV | WGT! |
| | 90 | 26 | 9G | 90 | 2 | 9 | 90 | 90 | PAATOS | AATOS | HG-PA | | | TEAID | TEATO | NEAID | KDVAD. | KEVAD. | NEAVD | ·VEATO | ·TEAD | TOKHED | DKHED | SSRLSD |
| 11 | KAPS | KAAS | KAPS | KAPS | KAPS | KAPS | KAPS | KAPS | ADVOF | VADVOF | WOTA | | 301 | 31 | DNFLFQWWDRLHK GTLTEAIDMKTI | DNFLFKWWDHVQK GTLNEAIDMKTI | DNFLFKWWDHVEE GKIKDVADMKTI | DNFLFKWWDHVEQ GIIKEVADMKTI | DNFLMGWWDHVEK GTLNEAVDMETI | DNFLMKWWDHVEK GTLVEAIDMKTI | DNFLMKAWDHVEK GTITEAADMKTI | ALN | ALN | RWRPG |
| 210 211 | , . } | | | | | | | -γ., | LHYC / | HYC | HYC | | 300 301 | HVQK (| REK (| HVQK (| HVEE (| HVEQ (| HVEK (| HVEK (| HVEK (| DLDA / | DLDA / | DLKR |
| | TPLGR | TPLGR | TTIGR | TPI GR | TPIGR | TPLGR | TPLGR | TPLGR | RFLTRI | RFI TRI | SFLTR | | | FKWD | FOWND | FKWD | FKWD | FKWD | MGWMD | MKMD | MKWD | ERWMA | ERWMA | PAWWG |
| 196 | VPVDQFTPLGRMLY- | VPVDEFTPLGRILY- | VPVDFFTTI GRALY- | VPVDFFTPI GRILY- | VPVDFFTPI GRILY- | VPVDOFTPLGRMLY- | VPVDEFTPLGRMLY- | VPVDEFTPLGRMLY- | PASAF | PASAF | PASSF | | 286 | NFL | NF | NFL | NF | NF | NF | NF | NF | RFL | IRFL | TOPAFL |
| 195 1 | - | - | | - | | | | | | | | | 285 2 | | | | | | | | | | | _ |
| | ELGIF | ELGIF |)FI G1F | FIGIE | F161F | ELGIF |)ELG1\ | ELGIV | 4F1 G1 | 11.014 | ELG19 | | | SPWF | SPWF | SPWFI | SPWFI | LSPWFI | LSPAFI | SPWFI | LSPAFI | ASPWF | ISPWF | WSPWF |
| | AGRKLLDELGIPAED | AORKL LDELGIPGAD | AORKI I DFI GIPAFD | AOBKI I DEI GIPAED | AORKI I DFI GIPAFD | AORKI LDELGIPAED | AORKLLDELG1VAED | AORKI FDELGIVAED | A I RKI FHEI G I PAHO | A TRKI FHEI GIPAHO | AVRKLQHELG1PPEQ | • | 71 | GEEGLKLSPWFRLVV | GEEGVKLSPWFKLIV | GEEGLKLSPWFRLVV | GEEGIKLSPWFRLVV | GEEGLKLSPWFRLVV | GEEGLKLSPWFRLVV | GEEGLKLSPWFRLVV | GDEAVKLSPWFRLVV | -DNGLQWSPWFRI1A | -DNGLQWSPWFRIIA | -DPGLSWSPWFRILA |
| ~ | ; ₩ | ¥ | ₹ | ν 2 | | | | | | | | | .7 | ច | 5 | 5 | _ | _ | _ | _ | _ | • | • | · |
| | | _ | ٠ ، | 5 | ting | } | na 1 | 2 00 | 115.1 | 11:00 | rts 2 rrdtii | | | | , | 2 | tina | ting | _ | ing 1 | ing 2 | lis 1 | ilis 2 | ırdtii |
| | 1 T.erecta | Sativa | 2 l cativa 2 | A A nalaestina | 2010 | מדומ | halia | halia | luvia | 7 | einho | | | recta | ativa | 3 L.sativa 2 | alaes | alaes | ativa | hal ia | halia: | luvia | luvia | reinha |
| | 1 7.6 | | יי היי | | י עיר עיר | 200 | 7 A. thaliana 1 | A A | 0 11 0 | 10 u nluvialis 2 | 10 n.piuviuiis 2 11 C.reinhardtii | | | 1 T.e | 2 L.sativa 1 | 3 6.5 | 4 A.D | 5 A.D | 6 <i>0.</i> s | 7 A.thaliana 1 | 8 A.t | 9 H.E | 10 H.pluvialis 2 | 11 C.reinhardtii 1 |
| | | | | | | | | | | _ | | • | | | | | | | | | | | | - |

F16.22B

+5 =



32 / 45

FIG. 24 A FIG. 23A FIG. 23 FIG. 24B FIG. 23B FIG. 24 F1G.23C FIG. 25A FIG.23D FIG. 25 FIG. 25B FIG. 25C

FIG.28A

FIG.28B

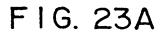
FIG. 28

10. 20

FIG. 26A

FIG. 26B

FIG. 26



| Comparison using GAP program of the Genetics Computer Group Gap Weight: 50 Average match: 10.000 Length Weight: 3 Average Mismatch: 0.000 Quality: 17392 Length: 1904 Ratio: 9.411 Gaps: 3 Percent Similarity: 95.331 Percent Identity: 95.331 Match display thresholds for the alignment(s): = IDENTITY := 5 . = 1 |
|--|
| Adonis palaestina ϵ -cyclase #3 x Adonis palaestina ϵ -cyclase #5 |
| 1 gagagaaaaagagtgttatattaatgttactgtcgcattcttgcaacac. 49 |
| 1aaaggagtgttctattaatgttactgtcgcattcttgcaacact 44 |
| 50 .atattcagactccattttcttgttttctcttcaaaacaacaactaatg 98 |
| 45 tatattcaaactccattttcttcttttctcttcaaaacaac |
| 99 tga.cggagtatctagctatggaactacttggtgttcgcaacctcatctc 147 |
| 95 tgagcagagtatétggétatggaactacttggtgttégéaacétéatété 144 |
| 148 ttcttgccctgtctggacttttggaacaagaaaccttagtagttcaaaac 197 |
| 145 ttettgeeetgtgtggaettttggaacaagaaacettagtagtteaaaac 194 |
| 198 tagcttataacatacatcgatatggttcttcttgtagagtagattttcaa 247 |
| 195 tágéttátáácátácátégátátágáttéttéttágtágágtág |
| 248 gtgagggctgatggtggaagcgggagtagaacttctgttgcttataaaga 297 |
| 245 gtgágagótgátgátgádágóggágtágáágttótgítgóttátádágá 294 |
| 298 gggttttgtggacgaggaggattttatcaaagctggtggttctgagcttt 347 |
| 295 gggttttgtggatgaagaggattttatcaaagctggtggttctgagcttt 344 |
| 348 tgtttgtccaaatgcagcaaacaaagtctatggagaaacaggccaagctc 397 |
| 345 tgtttgtccaaatgcagcaaacaaagtctatggagaaacaggccaagctc 394 |

FIG. 23B

| 398 | gccgataagttgccaccaatacctttcggagaatctgtgatggacttggt | 447 |
|-----|---|-----|
| | gccgataagttgccaccaataccttttggagaatccgtgatggacttggt | |
| 448 | tgtaataggttgtggacctgctggtctttcactggctgcagaagctgcta | 497 |
| 445 | tgtaataggttgtggacctgctggtctttcactggctgcagaagctgcta | 494 |
| 498 | agctaggcttgaaagttggccttattggtcctgatcttccttttacaaat | 547 |
| 495 | agctagggttgaaagttggccttattggtcctgatcttccttttacaaat | 544 |
| 548 | aattatggtgtgtgggaagacgagttcaaagatcttggacttgaacgttg | 597 |
| 545 | aattatggtgtgtgggaagacgagttcaaagatcttggacttgaacgttg | 594 |
| | tatcgagcatgcttggaaggacaccatcgtatatcttgacaatgatgctc | |
| | tatcgagcatgcttggaaggacaccatcgtatatcttgataatgatgctc | |
| | ctgtccttattggtcgtgcatatggacgagttagccggcatttgctgcat | 697 |
| | ctgtccttattggtcgtgcatatggacgagttagtcgacatttgctacat | 694 |
| | gaagagttgctgaaaaggtgtgtcgagtcaggtgtatcatatctgaattc | 747 |
| | gaggagttgctgaaaaggtgtgtggagtcaggtgtatcatatctggattc | |
| 745 | taaagtggaaaggatcactgaagctggtgatggccatagtcttgtagttt | |
| | taaagtggaaaggatcactgaagctggtgatggccatagccttgtagttt gtgaaaacgacatctttatcccttgcaggcttgctactgttgcatctgga | 794 |
| | | |
| | gcagcttcagggaaacttttggagtatgaagtaggtggccctcgtgtttg | |
| | | |
| | tgtccaaactgcttatggtgtggaggttgaggtggagaacaatccatacg | |
| 895 | tgtccaaaccgcttatggggtggaggttgaggtggagaacaatccatacg | |

FIG. 23C

| 948 a | atcccaacttaatggtatttatggactacagagactatatgcaacagaaa | 997 |
|--------|--|--------------|
| 945 a | atcccaacttaatggtattcatggactacagagactatatgcaacagaaa | 994 |
| 998 t | tacagtgctcggaagaagaatatccaacatttctctatgtcatgcccat | 1047 |
| 995 t | ttacagtgctcggaagaagaatatccaacatttctctatgtcatgcccat | 1044 |
| 1048 (| gtcgccaacaagacttttttttgaggaaacctgtttggcctcaaaagatg | 1097 |
| | gtcgccaacaagactttttttgaggaaacctgtttggcctcaaaagatg | 1094 |
| | ccatgcctttcgatctactgaagagaaaactaatgtcacgattgaagact | 1147 |
| | ccatgccattcgatctactgaagagaaaactgatgtcacgattgaagact | 1144 1197 |
| | ctgggtatccaagttacaaaaatttatgaagaggaatggtcttatattcc | 1197 |
| | tgttgggggttctttaccaaacacagagcaaaagaacctagcatttggtg | 1247 |
| | tgttggtggttctttaccaaacacagagcaaaagaacctagcatttggtg | 1244 |
| 1248 | ctgcagcaagcatggtgcatccagcaacaggctattcggttgtacgatca | 1297 |
| 1245 | ctgcagcaagcatggtgcatccagcaacaggctattcggttgtacggtca | 1294 |
| 1298 | ctatcagaagctccaaaatatgcttctgtaattgcaaagattttgaagca | 1347 |
| | ctgtcagaagctccaaaatatgcttctgtaattgcaaagattttgaagca | |
| 1348 | agataactctgcatatgtggtttctggacaaagcagtgcagtaaacattt | |
| | agataactctgcgtatgtggtttctggacaaagtagtgcagtaaacattt | _ |
| | caatgcaagcatggagcagtctttggccaaaggagcgaaaacgtcaaaga | |
| | caatgcaagcatggagcagtctttggccaaaggagcgaaaacgtcaaaga | ٠ |
| | gcattctttcttttcgggttagagcttattgtgcagctagatattgaagc | |
| 1445 | gcattettettttggattagagettattgtgeagetagatattgaage | . ITJ |

| 1498 | aaccagaacgttctttagaaccttcttccgcttgccaacttggatgtggt | 1547 |
|------|---|------|
| 1495 | aaccagaacattctttagaaccttcttccgcttgccaacttggatgtggt | 1544 |
| 1548 | ggggtttccttgggtcttcactatcatctttcgatcttgtattgtttcc | 1597 |
| 1545 | ggggtttccttgggtcttcactatcatctttcgatctcgtcttgttttcc | 1594 |
| 1598 | atgtacatgtttgttttggccccgaacagcatgaggatgtcacttgtgag | 1647 |
| 1595 | atgtacatgtttgttttggcgccaaacagcatgaggatgtcacttgtgag | 1644 |
| 1648 | acatttgctttcagatccttctggtgcagttatggttaaagcttacctcg | 1697 |
| 1645 | acatttgctttcagatccttctggtgcagttatggtaagagcttacctcg | 1694 |
| | aaaggtaatctgttttatgaaactatagtgtctcattaaataaatga | 1744 |
| | aaaggtagteteatetattattaaaetetagtgttteaceaaataaatga | 1744 |
| | <pre>ggatccttcgtatatgtatatgatcatctctatgtatatcctatattcta </pre> | 1794 |
| | ggatccttcgaatgtgtatatgatcatctctatgtatatcctgtactcta | 1794 |
| | atctcataaagtaatcgaaaattcattgatagaaaaaaaa | 1844 |
| | àtctcataaagtaaatgccgggtttgatattgttgtgtcaaaccggccaa | |
| | aaaa | 1848 |
| 1845 | tgatataaagtaaatttattgatacaaaagtagtttttttt | 1894 |

FIG. 23D

| • | | | - | | |
|----------------------------|---|---------------------------|--|--|---------------------------------------|
| GAP program blosum62.cm | of Genetics p | Computer | Group | FIG. | 24 A |
| Ga Lengt Percent Si | p Weight: h Weight: Quality: | ,147 ,623 s for the | Percent Ide alignment(| smatch: Length: Gaps: entity: (s): | 2.912 -2.003 530 0 99.057 |
| Adonis pala | estina ε-cyc | lase #3 | x Adonis pa | laestina & | E-cyclase #5 |
| | RNLİSSCPVWTF(| | | | |
| | VAYKEGFVDEEDI VAYKEGFVDEEDI | 11111 | | | |
| | VMDLVVIGCGPA VMDLVVIGCGPA | | | | |
| | .GLERCIEHAWKD GLERCIEHAWKD | | APVLIGRAYGRV APVLIGRAYGRV | | 200 |
| | SYLNSKVERITE . SYLDSKVERITE | | | | |
| | GPRVCVQTAYGV GPRVCVQTAYGV | | | | |
| | YVMPMSPTRLFF | | | | Ì |

| 351 | KIYEEEWSY] | IPVGGSLPNTEQKNLAFGAAASMVHPATGYSVVRSLSEAPK | 400 |
|-----|------------|--|-----|
| 351 | KVYEEEWSY] | | 400 |
| 401 | YASVIAKILK | KQDNSAYVVSGQSSAVNISMQAWSSLWPKERKRQRAFFLFG | 450 |
| 401 | YASVIAKILK | | 450 |
| 451 | LELIVOLDIE | EATRTFFRTFRLPTWMWWGFLGSSLSSFDLVLFSMYMFVL ! | 500 |
| 451 | LELIVOLDIE | ATRTFFRTFFRLPTWMWWGFLGSSLSSFDLVLFSMYMFVL | 500 |
| 501 | APNSMRMSLV | /RHLLSDPSGAVMVKAYLER* 530 | |
| 501 | APNSMRMSLV | | |

FIG. 24B



| | 103 102 102 107 107 108 72 73 73 73 | 54 205 204 204 209 202 202 191 181 182 182 182 185 |
|----------|---|--|
| 14 | | |
| | GPP GPP GPP GPP GPP GPP GPP GPP GPP GPP | 220 220 220 220 220 220 220 220 |
| | T. L. L. L. L. L. L. L. L. L. L. L. L. L. | PLLK PARCHO PARC |
| 100 | EGSK KQAK KQAK KQAK KQAK KQAK KPET VPET VPET | |
| | LECTOR COMP | SS SS HE SS SS SS SS SS SS SS SS SS SS SS SS SS |
| | TOON TOON TOON TOON TOON TOON TOON TOON | YGRV YGRV YGRV YGRV YGRV ARV |
| | FVQ FVQ FVQ FVQ VVSG VVKA VVKA VVKA VVKA VVKA VVKA VVKA | 200 16ka 16ka 16ka 16ka 16ka 16ka 16ka 16ka |
| | SSELI | DP1L KP1T KP1T APVL APVL DP1R PP1L TKO TKO TKO |
| 8 | KAGG KAGG KAGG KAGG IPC KAGG | POPULATION AND THE POPULATION AN |
| | EEDFU | TILLY TILLY AVENTY TILLY |
| • | 401 401 401 6 | 160 * 180 * 200 * |
| | KEGF KEGF (CGF) (CDF) (C | 180 AGIEHA AGIEH |
| | CVAV SVAY SVAY SVAY CVVDI CVVDI TSKL(SVKSV SVKSV SVKSV SVKSV SVKPP | |
| 90 | GSES GSRS GSRT GSRT SNNP SNNP SNNP SNNP SSFN ASTEI ASTEI ASTEI ASTEI ASTEI ASTEI ASTEI ASTEI ASTEI ASTEI ASTEI ASTEI | |
| | GGSS GGSS DGGS DGGS DGGS SAKSI NSSS NSSS -EKQI -VKV | |
| * | PASG QVRA QVRA QVRC GFE FG FG FG FG | NYGV NYGV NYGV NYGV NYGV |
| | SV RVDF RVDF RVDF RVDF RVDF PHHG PHHG PNHG PNHG | NEW NAME OF THE NA |
| | FPVVKRYSYRNIRFGL-CSVRASGGGSSGSESCVAVREDFADEEDFVKAGGSE1LFVQMQQNKDMDEQSKLVDKIJPPIS NLSSSKLAYNIHRYGSSCRVDFQVRADGGSGSRSSVAYKEGFVDEEDFIKAGGSELLFVQMQQTKSMEKQAKLADKIJPPIP HKFSLLKQRRFTNLSA-SSSLRQIKCSAKSDRCVVDKQISVADEEDFIKAGGSELLFVQMQQTKSMEKQAKLADKIJPPIP HKFSLLKQRRFTNLSA-SSSLRQIKCSAKSDRCVVDKQGISVADEEDFIKAGGSELLFVQMQQTKSMEKQAKLADKIJPPIP #SGGELCQEKSIFLAY-EQYESKCNSSSGSDSCVVDKEDFADEEDYIKAGGSELFVQMQQTKSMEKQAKLSDEIRQIS #SGGELCQEKSIFLAY-EQYEKQKONSSGSDSCVVDKEDFADEEDYIKAGGSQLVFVQMQQTKSMESQSKLSDEIRQISMOTLLKTPNKLDFFIPQFHGFERLCSNNPYHSRVRLGVKKRAIKIVSSVVSGSALLELVPETKKENIJDFELMOTLLRTPNNLEFLLPHHGFAEKQHLVSTSKLQNQVFRIASRNIHPCRNGTVKASSSALLELVPETKKENIJDFELMOTLLKTPNNLEFLHPVHGFSVKVSAFSSVVSGFGAKKFCEGLGSRSVCVKASSSALLELVPETKKENIJDFELMOTLLKTPNNLEFLHPVHGFSVKASSFNSVKPHKFGSRKICENWGKGVCVKAKSSALLELVPETKKENIJDFELMOTLLKTPNNLEFLLYPLHELAKRHFLSPSPNPQNPNFKFFSRKPPYQKKCRNGYIGVSSNQLLDLVPETKKENIJDFELMOTLLRTHNRLELLYPLHELAKRHFLSPSPNPQNPNFKFFSRKPPYQKKCRNGYIGVSSNQLLDLVPETKKENIJDFEL | 180 * 160 * 150 * 150 * 180 * 200 |
| 40 | WIRE ITHRY I | * |
| | CLAYN CGERRY CLERT CLEAT CLEAT CLEAT CLEAT CLEAT CLEAT | 40 SAKLGLKVGLLGPUP-AAKLGLKVGLLGPUP-SAKLGLKVGLLGPUP-SAKLGLKVGLLGPUP-SAKLGLKVGLLGPUP-SAKLGLKVGLVGPUP-VSEAGLSVCSTOPN-PKVSEAGLSVCSTOPN-PKVSEAGLSVVSTOPS-PKVSTOPS |
| * | SSSS SSSS SSSS SSSS SSSS SSSS SSSS SSSS SSSS | 40 SAKLG AAKLG AAKLG SAKLG SAKLG SAKLG VSEAG VSEAG |
| | STRNI STRNI STRNI STRNI | 14 AAES AAES AAES AAES AQQV AQQV AQQV AQQV |
| | \$ 5 5 5 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 | MGLAL MGLSL MGLSL GLAL GLAV GLAV GLAV GLAV GLAV |
| 20 | SCPW SCPW TCPR TCPR | * 25555555555555 |
| | MAVS - 1S: MAVE WARF | 120 * 140 * 160 * 160 * 160 * 180 * 200 * |
| * | F-AW 11ATI 11ATI | SOUVO DE LA COMPANION DE LA CO |
| | SARN SVRN SVRN SARN AG-H | 12 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 |
| A | **MECVGARNF-AAMAVSTFPSWS-CRRKFPVVKRYSYRNIRFGL-CSVRASGGGSSGSSCVAVREDFADEEDFVKAGGSELLFVQYQQTKSMECYGKLADKEPPIP **MELLGVRNLISSCPWIT-FGTRNLSSSKLAYNIHRYGSSCRVDFQVRAGGGSGRTSVAAYKEGFVDEEDFIKAGGSELLFVQYQQTKSMEKQAKLADKEPPIP **MELLGVRNLISSCPWIT-FGTRNLSSSKLAYNIHRYGSSCRVDFQVRAGGGSGRTSVAAYKEGFVDEEDFIKAGGSELLFVQYQQTKSMEKQAKLADKEPPIP **MECFGARNMTATMAVFTCPRFTDCNIRHKFSLLKQRRFTNLSA-SSSLRQIKCSAKSDRCVVDKQGISVADEEDFYKAGGSELFVQAQTKSMEXGRESQSKLSEKÄAQIP **MECFGARNMTATMAVFTCPRFTDCNIRHKFSLLKQRRFTNLSA-SSSLRQIKCSAKSDRCVVDKQGISVADEEDFYKAGGSELLFVQAQTKKDMDQQSKLSDELRQIS **MECFGARNMTATMAVFTCPRFM | 120 * 140 * 160 * 160 * 180 * 180 * 200 * |
| 25 | | ** ** ** ** ** ** ** ** ** ** ** ** |
| | PS is position of the position | psisl 1 2 2 3 3 5 8 8 8 |
| F16. 25A | ArabidopsisE ArabidopsisE AdonisE1 AdonisE2 TomatoE ArabidopsisB AdonisB PepperB TomatoB TomatoB MarigoldB | PotatoE ArabidopsisE AdonisE1 AdonisE2 LettuceEE TomatoE ArabidopsisB AdonisB PepperB TomatoB TomatoB OmatoB MarigoldB |
| غ بنا | Are Adc Adc Let Tom Mar Adc Pep Pep Tom Mar Daf | Pot Ara Ado Ado Iom Mar Pepi Tom Tom Mari |
| | | |

÷ ':

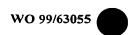


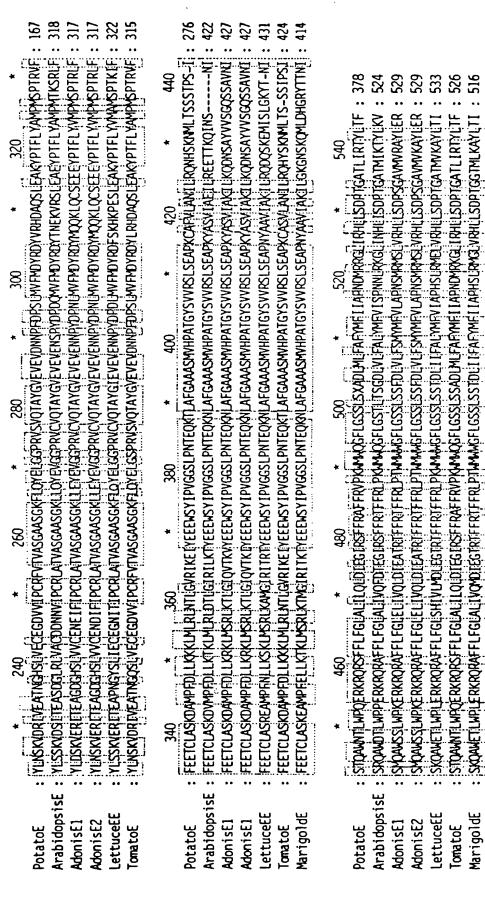
| | 4 | 10/45 |
|---|--|---|
| : 161 : 312 : 313 : 311 : 316 : 316 : 288 : 298 : 289 : 289 | 88,88 | . 271 |
| * 280 * 300 * 320 | GWKFHQAKÜIKVIH-EELKSLLICNDGVTIQATLVLDATG-FSRSLVQYD-KPYNPGYQVAYGILDEVEEHPFDVDKALFMDWRDSHLDQNLEIKARNSREPTFLYAMPF GWRFHQATÜVKAMH-EEEKSYLICSDGVTIDARVVLDATG-FSRCLVQYD-KPYNPGYQVAYGILDEVEEHPFDVDKAVFMDWRDSHLNGKAELNERNAKIPTFLYAMPF | 340 * 360 * 380 * 400 * 400 * 440 * |
| 3. -VRHDAQSLI -TNEKVRSLI -MQKLQCSI -MQKLQCSI -NGKLQCSI -SKHKPESLE | NNMELKERN QNLEIKARN GKAELNERN | O CÁFVLANŢÜ |
| MOYROY MOYROY MOYROY MOYROY MOYROY MOYROY MOYROY MOYROSHLK MOWROSHLK MOWROSHLK | MUWRUSHLG MUWRUSHLD MUWRUSHLN | VRSISEAPK |
| * 280 * 300 WYPPCRFVTVASGAASGKFLOYELGGPRVSVQTAYGVEVEVDNWIPEDPSLIW FINDYRDY INVEPCRLATVASGAASGKLLEYEVGGPRVCVQTAYGVEVEVENBYDPDQAYFINDYRDY IFIPCRLATVASGAASGKLLEYEVGGPRVCVQTAYGVEVEVENBYDPDLWAFINDYRDY IFIPCRLATVASGAASGKFLEYELGGPRVCVQTAYGVEVEVENBYDPDLWAFINDYRDY ITIPCRLATVASGAASGKFLEYELGGPRVCVQTAYGVEVEVENBYDPDLWAFINDYRDY ITIPCRLATVASGAASGKFLEYELGGPRVCVQTAYGVEVEVENBYDPDLWAFINDYRDY ITIPCRLATVASGAASGKFLQYELGGPRVCVQTAYGVEVEVEVENBYDPDLWAFINDYRDY ITIPCRLATVASGAASGKFLQYELGGPRVCVQTAYGTEVEVESTPYDPSLWFWINDWRDKHLI ITIPATVVLDATG-FSRCLVQYD-KPYNPGYQVAYGTWEVEHPFDVNKWYFMDWRDSHLN ITITQATVVLDATG-FSRSLVQYD-KPYNPGYQVAYGTLAEVEEHPFDVNKWYFMDWRDSHLN | PFDVDKYLF PFDVDKYLF PFDVDKYVF | * VHPATGÝSV |
| * *** *** *** *** *** *** *** *** *** | | 400 LAFGAGASM |
| 280 SPRVSVQTA SPRVCVQTA SPRVCVQTA SPRVSVQTA SPRVSVQTA SPRVSVQTA SPRVSVQTA SPRVSVQTA SPRVSVQTA SPRVSVQTA SPRVSVQTA SPRVSVQTA SPRVSVQTA SPRVSVQTA SPRVSVQTA SPRVSVQTA SPRVSVQTA SPRVSVQTA SPRVSVQVA SVRPGVQVA SVRPQVA SVRPQVQVA SVRPQVA | YNPGYQVAY YNPGYQVAY | * SERNTEOKT |
| GKFLQ [*] ELG GKLLQ [*] ELG GKLLQ [*] EVG GKLLEVEVG GKLLEVEVG GKLLQ [*] CLG GKLVQ [*] D-KF SCLVQ [*] D-KF SCLVQ [*] D-KF | SLVQYD-KP | 380 Ewsyi HV&G |
| 260 ATVASGAASI ATVASGA | 1081G-F3R | * WRIKEIYEE |
| * 240 * 300 | NDGVTÍDATLV SDGVTÍDARVV | 360 KLMLŘÚNTEG |
| * 240 GVLYLNSKÝDRIVEATNGHSLVEČEGD GVSYLSSKÝDS ITEASDGLRLVAČDDN GVSYLDSKÝERITEAGDGHSLVVČENE GVSYLNSKÝERITEAPNGYSLIEČEGN GVLYLNSKÝPRITEAPNGYSLIEČEGN GVKFHQSKÝTNVH-EEANSTVÝČSDGY GVKFHQAKÝTKVIH-EESKSLLIČNDGI GVKFHQAKÝTKVIH-EESKSMLIČNDGI GVKFHQAKÝTKVIH-EESKSMLIČNDGI GVKFHQAKÝTKVIH-EESKSMLIČNDGI GVKFHQAKÝTKVIH-EESKSMLIČNDGI GVKFHQAKÝTKVIH-EESKSMLIČNDGI | EELKSLLIĞ EEEKSYLIĞ | * DAMPFDLLKK |
| * VSKÝDRIVE SSKÝDSITE SSKÝERITE SSKÝERITE SSKÝERITE SSKÝTIVVH- AKVIKVIH- AKVIKVIH- AKVIKVIH- AKVIKVIH- | AKŲ IKVIH- ATŲ VKAMH- | 340 FEETCLASKI |
| | SVKFHQ SVRFHQ | SPTRVE |
| PotatoE ArabidopsisE AdonisE1 AdonisE2 LettuceEE TomatoE MarigoldE ArabidopsisB AdonisB TomatoB | MarigoldB DaffodilB | 340 * PotatoE : SPTRVFFEETGIASKDAMPFDLLKKKLM |

418 408 393 395 391 393 393 SLSEAPKCAFVLANŢĽRONHSKNMLTSS TKSRLFFEETICLASKOVMPFDLLKTKLMURUDTUGIRILKTYEEFJJSYIPYGGSLPNTEOKNLAFGAAASYVHPATGYSYYPRSLSEARKYASVIAEILREETTKQINS--SPTRUFFEETCLASKDAMPFDLLKRKLMSRLIKTLIGIQVTKVYFEEMSYIPVGGSLPNTEQKNLAFGAASMVHPATGYSVVRSLSEAPKYASVTAKTLKQDNSAYVVSGQ SPTRUFFEETCLASKDAMPFDLLKRKLMSRLIKTLIGIQVTKTYFEEMSYIPVGGSLPNTEQKNLAFGAAASMVHPATGYSVVRSLSEAPKYASVTAKTLKQDNSAYVVSGQ SPTKI FFEETCI JASREAMPFNLLKSKLMSRLIKAMGI RITRTYEEFJASYI PVGGSL PNTECKNLAFGAAASMVHPATGYSVVRSLSEAPNYAAVI AKTLRQDQSKEMI SLG SPTKVFFFETICI JASKEAMPFELLKTKLMSRLIKTMGIRITKTVEEFIASYI PVGGSL PINTEOKNLAFGAAJASMVHPATGYSVVRSI SEAPINYAAVIAK ILGKGNSKOMLDHG SPTRVFFEETCLASKDAMPFDLLKKKLMLRLINTLIGVRIKEIVEEELSYIHVGGSLANTEOKTLAFGAAASNVHPATGYSVVRSLSEAPKCASVLANILROHYSKNMLTS-SSNRIFILEETISI VARPGLRMEDIQERMAARIKHIGINVKRIGEDERCVII PYGGFLPVLPORVVGIGGTAGAVHPSTGYAVARTILAGAPIVANAIVRYLGSPSSN----S SSNRIFLEETSLVARPGLRIDDIQERMVÄRLINHLGIKVKSIEEDEHCLI PMGGPLPVLRORVGGGGTAGMVHPSTGYMÄRTLAAARVVANAI IQYLGSERSH-----SSNRIFLEETSLVARPGLRMDDIQERMVARLINHLGIKVKSIEEDEHCVI PMGGSLPVI PORVVGTGGTAGLVHPSTGYMVARTLAAARVVANAI IHYLGSEKOL-----SSTRIFLEETSLVARPGLKMEDIQERMAVRLKHLGIKVKSIEEDERCVI PMGGPLPVLRORVLGIGGTAGWVHPSTGYMVARTLAAARIVAKSI IRYLNNEKSM---VAD SSNRIFILETISLIVARPGLOMDDIQERAVARISHLIGIKVKSIEEDEHOVIPINGOPLAVLAQAVVGIGGIAGAVHASTGYAVARITLAAARIVVANAIIQYLSSERSH----SSTKIFLEETSLYARPGLRFEDIQERMARLKHLIGIKVKSI GEÖERCYI PYGGPL PVL PQRVVG IGGTAGYVHPSTGYMVARTILAAARVVAKSI VQYLGSDRSL SSNRIFIEETSLVARPGLIMEDIQEMVARLINHIGIRIKSIEFDERCVIPYGEPLEVIPORVOGIGGTAGMVHPSTGYMVA **SPTRVPPEETCJASKDAMPFOLLKKKLMĮRŲNTŪČVRIKE**TY ArabidopsisE **ArabidopsisB** MarigoldE ettuceEE Marigold8 Daffodi 1B AdonisE2 Adonis E1 FobaccoB AdonisB Pepper8 PotatoE TómatoB **Comato**E

FIG. 25B

| | # <u>`</u> | • | - |
|--|--|---|--|
| 41/45 | • | | |
| 378 378 378 378 378 378 378 378 | 3 a | : 50 : 208 : 207 : 207 : 212 | : 205 : 194 |
| * 'LER 'LER 'LER 'LER 'LER 'LOOKE | EPRVP EPRVP 220 | 7 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 | Sov |
| TLIRTY WWYRAY WWYKAY WWYKAY WWYKAY WWYKAY WWYKAY WWYKAY WWYKAY WWINNL WI | SISOK | | TRO S |
| 540 SUPTGATI | * * * | | |
| * INHLI: INHLI: VRHLL: VRHLL: WAK-G MAK-G MAK-G WAK-G | | GRVSR | GRVSRH |
| IDMRRGU ISMRASI ISMRASI ISLRAGU ITSRIEI ITSRIEI ITSRIEI ITSRIEI ITSRIEI ITSRIEI ITSRIEI ITSRIEI ITSRIEI ITSRIEI ITSRIEI ITSRIEI ITSRIEI | SELL FVC | HIGRAY TIGRAY UIGRAY UIGRAY RIGRAY | LIGRAY |
| 520 FILIAPN FI | TIKCNAAKSQLVVKQEIEEEEDYVKAGGSELLFVQVQQNKSYDAQSSISQKLPRVP | DODOKPI DNDAPV DNDAPV DDAUPI | DDOEP I |
| TEGES LA VACED STANDARD AND EED STANDARD | EEED | | TE LIVYLI JIVVYLI |
| SSADLY SSECTION * FLPELLY KEGF- * KEG | LVVKQEI 180 | TEHAWKI TEHAWK | |
| PLESSEL FLESSE | [V | | 01000 |
| SOO KANAGERI TANAGERI KANAGERI KANAGERI KANAGERI KANAGERI PRYAGERI PRYAGERI PRYAGERI PRYAGERI PRYAGERI PRYAGERI SOGSSSSS | CNAAKSC | | |
| * AFFRLP | KQIKCNAAKSQLVVKQEIEEEEDYVKAGGSELLFVQVQQNKSYDAQSSISQKLPRVP 160 * 180 * 220 | NNYGVH NNYGVH NNYGVH NNYGVH | NNYGVW |
| TRAFED TO SERVICE TO S | 1 | OLPFT OUPFT OUPFT | POLPFTI POLPFTI |
| STO STORY AND THE REPORT FOR THE TOTAL LUGIDIES FRAFER PROMAGE LESS IS SADIANT FOR THAN PRIBER AND STEPS LIST OF THE TOTAL PROPERTY OF THE TOTAL LUGIDIES FROM THE PROMAGE LESS IS SADIANT FOR THE TOTAL PROPERTY OF THE TOTAL LUGIDIES FROM THE PROMAGE LESS IS STORY THE TOTAL PROMAGE LIST FROM THE PROMAGE LESS IS STORY THE TOTAL PROMAGE LESS IS STORY THE PROMAGE LINE LESS CANNORANT LESS AND THE PROMAGE LINE LESS CANNORANT LESS AND THE PROMAGE LINE LESS CANNORANT LESS CANNORANT LINE LINE LINE LINE LINE LINE LINE LINE | TSIRYT | LGUKVGLIGPDLPFTNNYGVWEDE FKULGLQMCIEHYWRETIJYYLDDDKPITIGRAYGRYSRHLLHEELLKKYVFFSAYS LGUKVGLIGPDLPFTNNYGVWEDE FKOLGLERCIEHAWKOTIJYYLDNDAPVLIGRAYGRYSRHLLHEELLKRCVESGYS LGUKVGLIGPDLPFTNNYGVWEDE FKOLGLERCIEHAWKOTIJYYLDNDAPVLIGRAYGRYSRHLLHEELLKRCVESGYS LGUKVGLIGPDLPFTNNYGVWQDE FKOLGLERCIEHSWKOTIJYYLDDADPIRIGRAYGRYSRHLLHEELLKRCVESGYS | NGQTVI.DI.JVVIGCGPAGLALIARESAKI.GUNVGLVGPDLPFTNNYGVWEDEFKOLGUCACIEHVWROTI.IVVLDDOEPILIIGRAYGRVSRHFLHEELUKRCVENGV IGGGODSNCTIJDLVVIGCGPAGLALIAGESAKI.GUNVALIIGPDLPFTNNYGVWEDEFIGLGUEGCIEHVWROTIVVYLDDNOPILIIGRAYGRVSRDLLHEELUTROFESGVS |
| FEGLALIL FEG | 8 ! | SAKLGU | SAKLGU |
| Service of the servic | | SLAME SLAME SLAME MANE | 8 5 8 5 8 5 8 5 8 5 8 6 8 7 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 |
| STPS-ISTQAWITLWPGERGQRSFFLFGLA SSAVNISRQAWITLWPGERGQRSFFLFGLE SSAVNISRQAWFTLWPGERGQRSFFLFGLE SSAVNISRQAWFTLWPGERGQRSFFLFGLE KYT-NISRQAWFTLWPGERGQRSFFLFGLE RYTTNISRQAWFTLWPGERGQRSFFLFGLA RYTTNISRQAWFTLWPGERGQRSFFLFGLA SSIPSISTQAWNTLWPGERGAGRSFFLFGLA -SGNELSAEVWKDLWPTERRAGREFFCFGWD -SGNELSAEVWKDLWPTERRAGREFFCFGWD -SGNELSAAWWKDLWPTERRAGREFFCFGWD -SGNDLSAAWWKDLWPTERRAGREFFCFGWD | MSMRAG-HMTATMAAFTCPRFM- | IGDGALIDHVV IGCGPAGLALAAESAKI GESWICHVV IGCGPAGLSLAAEGAAKI GESWICHVV IGCGPAGLSLAAEGAKI IGNCTLICHVV IGCGPAGLALAAESAKI | QTVI:DI.VVIGCGPAGIALAGESAKI DSNC11:DI.VVIGCGPAGIALAGESAKI |
| ANTLWP SSSLWP COLW | TIMAF | | |
| * ISTOAN IISTO | (G-HMT) | | DSNC11 |
| FIG. 25C Potatoe : STPS- Adonise1 : SSAW Adonise1 : SSAW Adonise1 : SSAW Adonise1 : SSAW Arabidopsise : LRGD Arabidopsise : -SGNE FIG. 26A FIG. 26A FIG. 26A Arabidopsise : MECVE Arabidopsise : MECVE Arabidopsise : MECVE Arabidopsise : MECVE Arabidopsise : MECVE Adonise2 : MELLE Adonise2 : MELLE Adonise2 : MELLE Adonise2 : MELLE Adonise2 : MELLE Adonise2 : MELLE Adonise2 : MELLE Adonise2 : MELLE Adonise2 : MELLE Adonise2 : MELLE Adonise2 : MELLE | MSMRA | 4444 | A |
| | • •• | | •• |
| FG.2 Potatoe Arabidopsise Adonise1 Adonise1 Adonise Arabidopsise Arabidopsise Marigolde Marigolde Marigolde Marigolde Arabidopsise Adonise Potatoe Arabidopsise Adonise1 Adonise1 Adonise2 CettuceE | MarigoldE | Potatok ArabidopsisE AdonisE1 AdonisE2 LettuceEE | TomatoE MarigoldE |
| PCACOB Arabidopo AdonisE1 AdonisE1 CettuceEI TomatoE AdonisB AdonisB PepperB TomatoB TobaccoB MarigoldI Daffodill Daffodill Arabidopo AdonisE1 AdonisE2 LettuceEI TomatoF TomatoF | Maric | Potatoe Arabidop AdonisE1 AdonisE2 LettuceE | TomatoE Marigolo |
| | | | |





F16. 26B

43/45

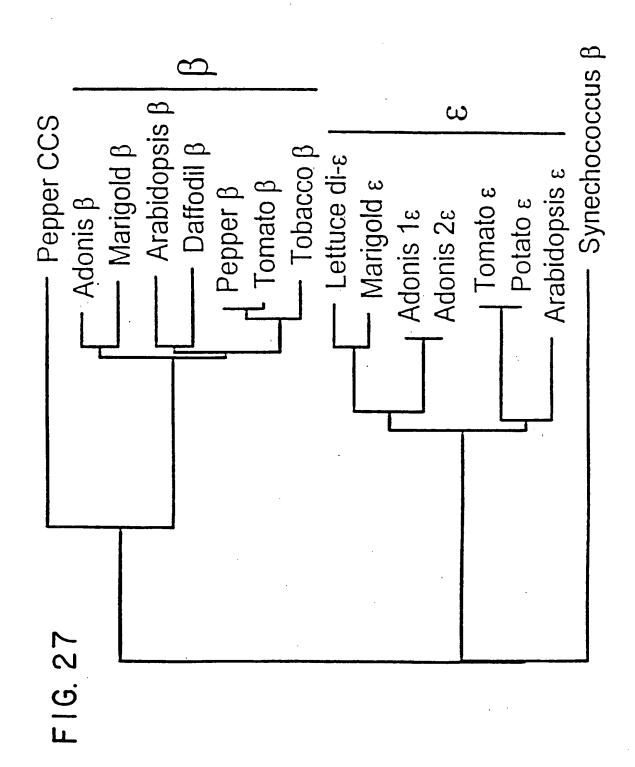


FIG. 28A

| Gap Weight: 12 Average Match: 2.912 Length Weight: 4 Average Mismatch: -2.003 Quality: 1837 Length: 534 Ratio: 3.499 Gaps: 3 Percent Similarity: 76.381 Percent Identity: 69.905 Match display thresholds for the alignment(s): = IDENTITY := 2 . = 1 |
|--|
| Arabidopsis x Lettuce |
| 1 MECVGARNF.AAMAVSTFPSWSCRRKFPVVKRYSYRNIRFGLCSVRA 46 |
| |
| 47 SGGGSSGSESCVAVREDFADEEDFVKAGGSEILFVQMQQNKDMDEQSKLV 96 |
| |
| 97 DKLPPISIGDGALDHVVIGCGPAGLALAAESAKLGLKVGLIGPDLPFTNN 146 |
| : . |
| 147 YGVWEDEFNDLGLQKCIEHVWRETIVYLDDDKPITIGRAYGRVSRRLLHE 196 |
| : : :::: |
| 197 ELLRRCVESGVSYLSSKVDSITEASDGLRLVACDDNNAIPCRLATVASGA 246 |
| 201 ELLRRCVESGVSYLSSKVERITEAPNGYSLIECEGNITIPCRLATVASGA 250 |
| 247 ASGKLLQYEVGGPRVCVQTAYGVEVEVENSPYDPDQMVFMDYRDYTNEKV 296 |
| 251 ASGKFLEYELGGPRVCVQTAYGIEVEVENNPYDPDLMVFMDYRDFSKHKP 300 |
| 297 RSLEAEYPTFLYAMPMTKSRLFFEETCLASKDVMPFDLLKTKLMLRLDTL 346 |
| . : :: :: . . : 301 ESLEAKYPTFLYVMAMSPTKIFFEETCLASREAMPFNLLKSKLMSRLKAM 350 |

FIG. 28B

| | • | |
|-----|--|-----|
| 347 | GIRILKTYEEEWSYIPVGGSLPNTEQKNLAFGAAASMVHPATGYSVVRSL | 396 |
| 351 | : | 400 |
| 397 | SEAPKYASVÍAEILREETTKQINSNISRQAWDTLWPPERKRQRAF | 441 |
| 401 | . . :: . :. : : SEAPNYAAVIAKILRQDQSKEMISLGKYTNISKQAWETLWPLERKRQRAF | 450 |
| 442 | FLFGLALIVOFDTEGIRSFFRTFFRLPKWMWQGFLGSTLTSGDLVLFALY | 491 |
| 451 | . | 500 |
| 492 | MFVISPNNLRKGLINHLISDPTGATMIKTYLKV* 525 | |
| 501 | . : : : : MFVIAPHSLRMELVRHLLSDPTGATMVKAYLTI* 534 | |



SEQUENCE LISTING

| <110> | CUNN: | INGHA ZAIF | AM JE REN | Ŗ., F | RANC | :IS } | ۲. | | | | | | | | |
|---------------------------|----------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|-----|
| <120> | GENE: | S OF ODS C | CARC F US | TENC | ID E | SIOSY F | NTHE | ESIS | AND | META | ABOLI | ISM A | AND | | |
| <130> | 8172- | -9023 | 3 | | | | | | | | | | | | |
| <140> <141> | NOT 1 | YET A | \SSIG | NED | | | | | | | | | | | |
| <150> <151> | 09/08 1998- | 38,72 -06-0 | 2 | | | | | | | | | | | | |
| <150> <151> | 09/08 1998- | 88,72 -06-0 | 5 | | | | | | | | | | | | |
| <160> | 61 | | | | | | | | | | | | | | |
| <170> | Pater | tIn | Ver. | 2.0 | | | | | | | | | | | |
| <210><211><211><212><213> | 1860 DNA | dops | is t | hali | ana | | | | | | | | | | |
| <220> <221> <222> | | (1 | 680) | | | | | | | | | | | | |
| <400> ACAAAA | _ | ATAA | TTAG. | AT T | CCTC' | TTTC | T GC | TTGC | ТАТА | ССТ | TGAT | AGA . | ACAA | ТАТААС | 60 |
| AATGGT | 'GTAA | GTCT | TCTC | GC T | GTAT' | TCGA | А АТ' | TATT | TGGA | GGA | GGAA | Me | | G TGT u Cys | 117 |
| GTT GG Val Gl | G GCT y Ala 5 | AGG Arg | AAT Asn | TTC Phe | GCA Ala 10 | GCA Ala | ATG Met | GCG Ala | GTT Val | TCA Ser 15 | ACA Thr | TTT Phe | CCG Pro | TCA Ser | 165 |
| TGG AG Trp Se 20 | T TGT r Cys | CGA Arg | AGG Arg | AAA Lys 25 | TTT Phe | CCA Pro | GTG Val | GTT Val | AAG Lys 30 | AGA Arg | TAC Tyr | AGC Ser | TAT Tyr | AGG Arg 35 | 213 |
| AAT AT Asn Il | T CGT e Arg | TTC Phe | GGT Gly 40 | TTG Leu | TGT Cys | AGT Ser | GTC Val | AGA Arg 45 | GCT Ala | AGC Ser | GGC Gly | GGC Gly | GGA Gly 50 | AGT Ser | 261 |
| TCC GG Ser Gl | T AGT y Ser | GAG Glu 55 | AGT Ser | TGT Cys | GTA Val | GCG Ala | GTG Val 60 | AGA Arg | GAA Glu | GAT Asp | TTC Phe | GCT Ala 65 | GAC Asp | GAA Glu | 309 |
| GAA GA Glu As | T TTT p Phe 70 | GTG Val | AAA Lys | GCT Ala | GGT Gly | GGT Gly 75 | TCT Ser | GAG Glu | ATT Ile | CTA Leu | TTT Phe 80 | GTT Val | CAA Gln | ATG Met | 357 |
| CAG CA | G AAC | AAA | GAT | ATG | GAT | GAA | CAG | TCT | AAG | CTT | GTT | GAT | AAC | ም ምር | 405 |

PCT/US99/12121





| Gln | Gln 85 | Asn | Lys | Asp | Met | Asp 90 | Glu | Gln | Ser | Lys | Leu 95 | Val | Asp | Lys | Leu | |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|------|
| CCT Pro 100 | CCT Pro | ATA Ile | TCA Ser | ATT Ile | GGT Gly 105 | GAT Asp | GGT Gly | GCT Ala | TTG Leu | GAT Asp 110 | CAT | GTG Val | GTT Val | ATT Ile | GGT Gly 115 | 453 |
| TGT Cys | GGT Gly | CCT Pro | GCT Ala | GGT Gly 120 | Leu | GCC Ala | TTG Leu | GCT Ala | GCA Ala 125 | GAA Glu | TCA Ser | GCT Ala | AAG Lys | CTT Leu 130 | GGA Gly | 501 |
| TTA Leu | AAA Lys | GTT Val | GGA Gly 135 | CTC Leu | ATT Ile | GGT Gly | CCA Pro | GAT Asp 140 | CTT Leu | CCT Pro | TTT Phe | ACT Thr | AAC Asn 145 | AAT Asn | TAC Tyr | 549 |
| GGT Gly | GTT Val | TGG Trp 150 | GAA Glu | GAT Asp | GAA Glu | TTC Phe | AAT Asn 155 | GAT Asp | CTT Leu | GGG Gly | CTG Leu | CAA Gln 160 | AAA Lys | TGT Cys | ATT Ile | 597 |
| GAG Glu | CAT His 165 | GTT Val | TGG Trp | AGA Arg | GAG Glu | ACT Thr 170 | ATT Ile | GTG Val | TAT Tyr | CTG Leu | GAT Asp 175 | GAT Asp | GAC Asp | AAG Lys | CCT Pro | 645 |
| | | | | | | | | | | | | | TTG Leu | | | 693 |
| | | | | | | | | | | | | | TAC Tyr | | | 741 |
| TCG Ser | AAA Lys | GTT .Val | GAC Asp 215 | AGC Ser | ATA Ile | ACA Thr | GAA Glu | GCT Ala 220 | TCT Ser | GAT Asp | GGC Gly | CTT Leu | AGA Arg 225 | CTT Leu | GTT Val | 789 |
| | | | | | | | | | | | | | ACT Thr | | | 837 |
| | | | | | | | | | | | | | GGT Gly | | | 885 |
| | | | | | | | | | | | | | GTG Val | | | 933 |
| AGT Ser | CCA Pro | TAT Tyr | GAT Asp | CCA Pro 280 | GAT Asp | CAA Gln | ATG Met | GTT Val | TTC Phe 285 | ATG Met | GAT Asp | TAC Tyr | AGA Arg | GAT Asp 290 | TAT Tyr | 981 |
| | | | | | | | | | | | | | ACG Thr 305 | | | 1029 |
| | | | | | | | | | | | | | GAG Glu | | | 1077 |
| | | | | | | | | | | | | | ACG Thr | | | 1125 |

| WO | 99/63055 | , |
|----|----------|---|

PCT/US99/12121

| 325 | . 33 | 0 | 335 | e j | |
|---|---|---------------------------------------|---|---------------------------------|------|
| ATG TTA AGA T Met Leu Arg Lo 340 | TA GAT ACA CTO eu Asp Thr Le 345 | C GGA ATT CGA u Gly Ile Arg | ATT CTA AAG A Ile Leu Lys T 350 | CT TAC GAA hr Tyr Glu 355 | 1173 |
| GAG GAG TGG TG Glu Glu Trp Se | CC TAT ATC CC er Tyr Ile Pro 360 | A GTT GGT GGT > Val Gly Gly 365 | TCC TTG CCA A Ser Leu Pro A | AC ACC GAA sn Thr Glu 370 | 1221 |
| Gin Lys Asn Le | rc GCC TTT GG eu Ala Phe Gl 75 | r GCT GCC GCT y Ala Ala Ala 380 | AGC ATG GTA CASER Met Val H. | AT CCC GCA is Pro Ala 85 | 1269 |
| ACA GGC TAT TO Thr Gly Tyr Se 390 | CA GTT GTG AGA er Val Val Arc | A TCT TTG TCT g Ser Leu Ser 395 | GAA GCT CCA AG Glu Ala Pro L 400 | AA TAT GCA ys Tyr Ala | 1317 |
| TCA GTC ATC GG Ser Val Ile AI 405 | CA GAG ATA CTA La Glu Ile Lev 410 | ı Arg Glu Glu | ACT ACC AAA CA Thr Thr Lys GO 415 | AG ATC AAC ln Ile Asn | 1365 |
| AGT AAT ATT TO Ser Asn Ile Se 420 | CA AGA CAA GC er Arg Gln Ala 425 | T TGG GAT ACT Trp Asp Thr | TTA TGG CCA CC Leu Trp Pro Pr 430 | CA GAA AGG ro Glu Arg 435 | 1413 |
| AAA AGA CAG AG Lys Arg Gln Ar | GA GCA TTC TT rg Ala Phe Phe 440 | CTC TTT GGT Leu Phe Gly 445 | CTT GCA CTC AT Leu Ala Leu II | FA GTT CAA le Val Gln 450 | 1461 |
| TTC GAT ACC GAT Phe Asp Thr G1 | lu Gly Ile Arq | A AGC TTC TTC g Ser Phe Phe 460 | Arg Thr Phe Ph | TC CGC CTT ne Arg Leu 65 | 1509 |
| CCA AAA TGG AT Pro Lys Trp Me 470 | CG TGG CAA GGG | G TTT CTA GGA 7 Phe Leu Gly 475 | TCA ACA TTA AC Ser Thr Leu Th 480 | CA TCA GGA or Ser Gly | 1557 |
| GAT CTC GTT CT Asp Leu Val Le 485 | CC TTT GCT TTA eu Phe Ala Leu 490 | Tyr Met Phe | GTC ATT TCA CO Val Ile Ser Pr 495 | CA AAC AAT | 1605 |
| TTG AGA AAA GG Leu Arg Lys Gl 500 | GT CTC ATC AAT y Leu Ile Asr 505 | His Leu Ile | TCT GAT CCA AC Ser Asp Pro Th 510 | CC GGA GCA or Gly Ala 515 | 1653 |
| ACC ATG ATA AF Thr Met Ile Ly | AA ACC TAT CTO s Thr Tyr Leu 520 | C AAA GTA TGAT 1 Lys Val | TTACTT ATCAACT | TCTT : | 1700 |
| AGGTTTGTGT ATA | TATATGT TGAT | TTATCT GAATAAT | CGA TCAAAGAATO | G GTATGTGGGT | 1760 |
| TACTAGGAAG TTG | GAAACAA ACATO | STATAG AATCTAA | AGGA GTGATCGAAA | A TGGAGATGGA | 1820 |
| AACGAAAAGA AAA | AAATCAG TCTT | GTTTT GTGGTTA | AGTG | : | 1860 |

<210> 2 <211> 524 <212> PRT <213> Arabidopsis thaliana



<400> 2 Met Glu Cys Val Gly Ala Arg Asn Phe Ala Ala Met Ala Val Ser Thr Phe Pro Ser Trp Ser Cys Arg Arg Lys Phe Pro Val Val Lys Arg Tyr Ser Tyr Arg Asn Ile Arg Phe Gly Leu Cys Ser Val Arg Ala Ser Gly Gly Gly Ser Ser Gly Ser Glu Ser Cys Val Ala Val Arg Glu Asp Phe Ala Asp Glu Glu Asp Phe Val Lys Ala Gly Gly Ser Glu Ile Leu Phe Val Gln Met Gln Gln Asn Lys Asp Met Asp Glu Gln Ser Lys Leu Val Asp Lys Leu Pro Pro Ile Ser Ile Gly Asp Gly Ala Leu Asp His Val 100 Val Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala Ala Glu Ser Ala Lys Leu Gly Leu Lys Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Glu Asp Glu Phe Asn Asp Leu Gly Leu Gln Lys Cys Ile Glu His Val Trp Arg Glu Thr Ile Val Tyr Leu Asp Asp Asp Lys Pro Ile Thr Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg Arg Leu Leu His Glu Glu Leu Leu Arg Arg Cys Val Glu Ser Gly Val Ser 200 Tyr Leu Ser Ser Lys Val Asp Ser Ile Thr Glu Ala Ser Asp Gly Leu Arg Leu Val Ala Cys Asp Asp Asn Asn Val Ile Pro Cys Arg Leu Ala Thr Val Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Gln Tyr Glu Val Gly Gly Pro Arg Val Cys Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Glu Asn Ser Pro Tyr Asp Pro Asp Gln Met Val Phe Met Asp Tyr 280 Arg Asp Tyr Thr Asn Glu Lys Val Arg Ser Leu Glu Ala Glu Tyr Pro Thr Phe Leu Tyr Ala Met Pro Met Thr Lys Ser Arg Leu Phe Phe Glu 315

WO 99/63055



360

420

480

| Glu | Thr | Cys | Leu | Ala 325 | Ser | Lys | Asp | Val | Met 330 | Pro | Phe | Asp | Leu | Leu 335 | Lys | |
|------------|------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-----|
| Thr | Lys | Leu | Met 340 | Leu | Arg | Leu | Asp | Thr 345 | Leu | Gly | Ile | Arg | Ile 350 | Leu | Lys | |
| Thr | Tyr | Glu 355 | Glu | Glu | Trp | Ser | Tyr 360 | Ile | Pro | Val | Gly | Gly 365 | Ser | Leu | Pro | |
| Asn | Thr 370 | Glu | Gln | Lys | Asn | Leu 375 | Ala | Phe | Gly | Ala | Ala 380 | Ala | Ser | Met | Val | |
| His 385 | Pro | Ala | Thr | Gly | Tyr 390 | Ser | Val | Val | Arg | Ser 395 | Leu | Ser | Glu | Ala | Pro 400 | |
| Lys | Tyr | Ala | Ser | Val 405 | Ile | Ala | Glu | Ile | Leu 410 | Arg | Glu | Glu | Thr | Thr 415 | Lys | |
| Gln | Ile | Asn | Ser 420 | Asn | Ile | Ser | Arg | Gln 425 | Ala | Trp | Asp | Thr | Leu 430 | Trp | Pro | |
| Pro | Glu | Arg 435 | Lys | Arg | Gln | Arg | Ala 440 | Phe | Phe | Leu | Phe | Gly 445 | Leu | Ala | Leu | |
| Ile | Val 450 | Gln | Phe | Asp | Thr | Glu 455 | Gly | Ile | Arg | Ser | Phe 460 | Phe | Arg | Thr | Phe | |
| Phe 465 | Arg | Leu | Pro | Lys | Trp 470 | Met | Trp | Gln | Gly | Phe 475 | Leu | Gly | Ser | Thr | Leu 480 | |
| Thr | Ser | Gly | Asp | Leu 485 | Val | Leu | Phe | Ala | Leu 490 | Tyr | Met | Phe | Val | Ile 495 | Ser | |
| Pro | Asn | Asn | Leu 500 | Arg | Lys | Gly | Leu | Ile 505 | Asn | His | Leu | Ile | Ser 510 | Asp | Pro | |
| Thr | Gly | Ala 515 | Thr | Met | Ile | Lys | Thr 520 | Tyr | Leu | Lys | Val | | | | | |
| <212 | l > 95 2 > DN | ΙA | dopsi | is th | nalia | na | | | | | | | | | | |
| |)> 3 C T TTC | CTC (| CTCCI | CCT | CT AC | CCGAT | TTTC | C GAC | CTCCC | SCCT | CCC | TAAA | rcc 1 | TATO | CCGGAT | 60 |
| | | | | | | | | | | | | | | | GAGAC | 120 |
| | | | | | | | | | | | | | | | CTATAG | 180 |
| | | | | | | | | | | | | | | | AGAGGT | 240 |
| | | | | | | | | | | | | | | | TATGG | 300 |

CTGTTTACTA CAGATTCTCT TGGCAAATGG AGGGAGGTGA GATCTCAATG TTGGAAATGT

TTGGTACATT TGCTCTCTC GTTGGTGCTG CTGTTGGTAT GGAATTCTGG GCAAGATGGG

CTCATAGAGC TCTGTGGCAC GCTTCTCTAT GGAATATGCA TGAGTCACAT CACAAACCAA

WO 99/63055

| \ | | | |
|----------|--|--|--|
| • | | | |
| • | | | |

PCT/US99/12121

| | <i>'</i> | - | | | | | | | | | |
|--|----------------|----------------|------------|------------|------------|------------|------------|------------|------------|------------|----------------------|
| GAGAAGGACC | GTTTGAGC | TA AACGA | TGTT | r ttc | SCTAT | TAGT | GAA | CGCT | GGT (| CCAG | CGATTG |
| GTCTCCTCTC | TTATGGAT | тс ттсаа | TAAAC | G GAC | CTCGT | TCC | TGG' | rctc: | rgc ' | TTTG | GCGCC ['] G |
| GGTTAGGCAT | AACGGTGT' | TT GGAAT | CGCCI | r aca | ATGTT | TGT | CCA | CGAT | GGT (| CTCG' | rgcaca |
| AGCGTTTCCC | TGTAGGTC | CC ATCGC | CGAC | G TCC | CTTA | ACCT | ÇCG | AAAG | STC (| GCCG | CCGCTC |
| ACCAGCTACA | TCACACAG | AC AAGTT | CAATO | G GTC | STACC | CATA | TGG | ACTG | TTT (| CTTG | GACCCA |
| AGGAATTGGA | AGAAGTTG | GA GGAAA | TGAAC | G AGT | TAGA | AATA | GGA | SATTA | AGT (| CGGA | SAATCA |
| AATCATACAA | AAAGGCCT | CG GGCTC | CGGGT | r CGA | AGTTC | CGAG | TTC | TGAG | CTT |)AAA1 | CAAGTT |
| TTAAATCCCA | AATTCTTT' | TT TTGTC | TTCT | TCF | LATTA | GAT | CATO | CTTA | AGA (| CGGT | CT |
| <210> 4 <211> 294 <212> PRT <213> Arabi | dopsis t | naliana | | | | | | | | | |
| <400> 4 Ser Phe Ser | Ser Ser | Ser Thr | Asp | Phe | Ara | Leu | Ara | Len | Pro | Lus | Ser |
| 1 . | 5 | | | | 10 | | | | | 15 | |
| Leu Ser Gly | Phe Ser 20 | Pro Ser | Leu | Arg 25 | Phe | Lys | Arg | Phe | Ser 30 | Val | Cys |
| Tyr Val Val 35 | Glu Glu | Arg Arg | Gln 40 | Asn | Ser | Pro | Ile | Glu 45 | Asn | Asp | Glu |
| Arg Pro Glu 50 | Ser Thr | Ser Ser 55 | Thr | Asn | Ala | Ile | Asp 60 | Ala | Glu | Tyr | Leu |
| Ala Leu Arg 65 | Leu Ala | Glu Lys 70 | Leu | Glu | Arg | Lys 75 | Lys | Ser | Glu | Arg | Ser 80 |
| Thr Tyr Leu | Ile Ala 85 | Ala Met | Leu | Ser | Ser 90 | Phe | Gly | Ile | Thr | Ser 95 | Met. |
| Ala Val Met | Ala Val 100 | Tyr Tyr | Arg | Phe 105 | Ser | Trp | Gln | Met | Glu 110 | Gly | Gly |
| Glu Ile Ser 115 | Met Leu | Glu Met | Phe 120 | Gly | Thr | Phe | Ala | Leu 125 | Ser | Val | Gly |
| Ala Ala Val 130 | Gly Met | Glu Phe 135 | Trp | Ala | Arg | Trp | Ala 140 | His | Arg | Ala | Leu |
| Trp His Ala 145 | Ser Leu | Trp Met 150 | Asn | His | Glu | Ser 155 | His | His | Lys | Pro | Arg 160 |
| Glu Gly Pro | Phe Glu 165 | Leu Asn | Asp | Val | Phe 170 | Ala | Ile | Val | Asn | Ala 175 | Gly |
| Pro Ala Ile | Gly Leu 180 | Leu Ser | Tyr | Gly 185 | Phe | Phe | Asn | Lys | Gly 190 | Leu | Val |
| Pro Gly Leu 195 | Cys Phe | Gly Ala | Gly 200 | Leu | Gly | Ile | Thr | Val 205 | Phe | Gly | Ile |



- Ala Tyr Met Phe Val His Asp Gly Leu Val His Lys Arg Phe Pro Val 210 220
- Gly Pro Ile Ala Asp Val Pro Tyr Leu Arg Lys Val Ala Ala Ala His 225 230 235 240
- Gln Leu His His Thr Asp Lys Phe Asn Gly Val Pro Tyr Gly Leu Phe 245 250 255
- Leu Gly Pro Lys Glu Leu Glu Glu Val Gly Gly Asn Glu Glu Leu Asp 260 265 270
- Lys Glu Ile Ser Arg Arg Ile Lys Ser Tyr Lys Lys Ala Ser Gly Ser 275 280 285
- Gly Ser Ser Ser Ser Ser 290
- <210> 5
- <211> 162
- <212> PRT
- <213> Alicalgenes sp.
- <400> 5
- Met Thr Gln Phe Leu Ile Val Val Ala Thr Val Leu Val Met Glu Leu 1 5 10 15
- Thr Ala Tyr Ser Val His Arg Trp Ile Met His Gly Pro Leu Gly Trp 20 25 30
- Gly Trp His Lys Ser His His Glu Glu His Asp His Ala Leu Glu Lys
 35 40 45
- Asn Asp Leu Tyr Gly Val Val Phe Ala Val Leu Ala Thr Ile Leu Phe 50 55 60
- Thr Val Gly Ala Tyr Trp Trp Pro Val Leu Trp Trp Ile Ala Leu Gly 65 70 75 80
- Met Thr Val Tyr Gly Leu Ile Tyr Phe Ile Leu His Asp Gly Leu Val 85 90 95
- His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Arg Gly Tyr Phe Arg 100 105 110
- Arg Leu Tyr Gln Ala His Arg Leu His His Ala Val Glu Gly Arg Asp 115 120 125
- His Cys Val Ser Phe Gly Phe Ile Tyr Ala Pro Pro Val Asp Lys Leu 130 135 140
- Lys Gln Asp Leu Lys Arg Ser Gly Val Leu Arg Pro Gln Asp Glu Arg 145 150 155 160
- Pro Ser
- <210> 6
- <211> 175
- <212> PRT
- <213> Erwinia herbicola



Ile Ala Ala Phe Thr His Arg Tyr Ile Met His Gly Trp Gly Trp Arg
20 .25 30

Trp His Glu Ser His His Thr Pro Arg Lys Gly Val Phe Glu Leu Asn 35 40 45

Asp Leu Phe Ala Val Val Phe Ala Gly Val Ala Ile Ala Leu Ile Ala 50 55 60

Val Gly Thr Ala Gly Val Trp Pro Leu Gln Trp Ile Gly Cys Gly Met 65 70 75 80

Thr Val Tyr Gly Leu Leu Tyr Phe Leu Val His Asp Gly Leu Val His 85 90 95

Gln Arg Trp Pro Phe His Trp Ile Pro Arg Arg Gly Tyr Leu Lys Arg 100 105 110

Leu Tyr Val Ala His Arg Leu His His Ala Val Arg Gly Arg Glu Gly 115 120 125

Cys Val Ser Phe Gly Phe Ile Tyr Ala Arg Lys Pro Ala Asp Leu Gln 130 135 140

Ala Ile Leu Arg Glu Arg His Gly Arg Pro Pro Lys Arg Asp Ala Ala 145 150 155 160

Lys Asp Arg Pro Asp Ala Ala Ser Pro Ser Ser Ser Pro Glu 165 170 175

<210> 7

<211> 175

<212> PRT

<213> Erwinia uredovora

<400> 7

Met Leu Trp Ile Trp Asn Ala Leu Ile Val Phe Val Thr Val Ile Gly
1 5 10 15

Met Glu Val Ile Ala Ala Leu Ala His Lys Tyr Ile Met His Gly Trp 20 25 30

Gly Trp Gly Trp His Leu Ser His His Glu Pro Arg Lys Gly Ala Phe 35 40

Glu Val Asn Asp Leu Tyr Ala Val Val Phe Ala Ala Leu Ser Ile Leu 50 55 60

Leu Ile Tyr Leu Gly Ser Thr Gly Met Trp Pro Leu Gln Trp Ile Gly 65 70 75 80

Ala Gly Met Thr Ala Tyr Gly Leu Leu Tyr Phe Met Val His Asp Gly 85 90 95

Leu Val His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr 100 105 110

WO 99/63055



PCT/US99/12121

| | | | | | | | | | | | | | - | | | |
|---|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-----|
| Leu | Lys | Arg 115 | Leu | Tyr | Met | Ala | His 120 | Arg | Met | His | His | Ala 125 | Val | Arg | Gly | |
| Lys | Glu 130 | Gly | Cys | Val | Ser | Phe 135 | Gly | Phe | Leu | Tyr | Ala 140 | Pro | Pro | Leu | Ser | |
| Lys 145 | Leu | Gln | Ala | Thr | Leu 150 | Arg | Glu | Arg | His | Gly 155 | Ala | Arg | Ala | Gly | Ala 160 | |
| Ala | Arg | Asp | Ala | Gln 165 | Gly | Gly | Glu | Asp | Glu 170 | Pro | Ala | Ser | Gly | Lys 175 | | |
| <210> 8 <211> 162 <212> PRT <213> Agrobacterium aurianticum | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | |
| | 0> 8 Thr | Asn | Phe | Leu 5 | Ile | Val | Val | Ala | Thr 10 | Val | Leu | Val | Met | Glu 15 | Leu | |
| Thr | Ala | Tyr | Ser 20 | Val | His | Arg | Trp | Ile 25 | Met | His | Gly | Pro | Leu 30 | Gly | Trp | |
| Gly | Trp | His 35 | Lys | Ser | His | His | Glu 40 | Glu | His | Asp | His | Ala 45 | Leu | Glu | Lys | |
| Asn | Asp 50 | Leu | Tyr | Gly | Leu | Val 55 | Phe | Ala | Val | Ile | Ala 60 | Thr | Val | Leu | Phe | |
| Thr 65 | Val | Gly | Trp | Ile | Trp 70 | Ala | Pro | Val | Leu | Trp 75 | Trp | Ile | Ala | Leu | Gly 80 | |
| Met | Thr | Val | Tyr | Gly 85 | Leu | Ile | Tyr | Phe | Val 90 | Leu | His | Asp | Gly | Leu 95 | Val | |
| His | Trp | Arg | Trp 100 | Pro | Phe | Arg | Tyr | Ile 105 | Pro | Arg | Lys | Gly | Tyr 110 | Ala | Arg | |
| Arg | Leu | Tyr 115 | Gln | Ala | His | Arg | Leu 120 | His | His | Ala | Val | Glu 125 | Gly | Arg | Asp | |
| His | Cys 130 | Val | Ser | Phe | Gly | Phe 135 | Ile | Tyr | Ala | Pro | Pro 140 | Val | Asp | Lys | Leu | |
| Lys 145 | Gln | Asp | Leu | Lys | Met 150 | Ser | Gly | Val | Leu | Arg 155 | Ala | Glu | Ala | Gln | Glu 160 | |
| Arg | Thr | | | | | | | | | | | | | | | |
| <210> 9 <211> 954 <212> DNA <213> Arabidopsis thaliana | | | | | | | | | | | | | | | | |
| <400> 9 CCACGGGTCC GCCTCCCCGT TTTTTTCCGA TCCGATCTCC GGTGCCGAGG ACTCAGCTGT 60 | | | | | | | | | | | | | | | | |
| TTGT | TCGC | GC 1 | TTCI | CAGO | CC G1 | CAC | CATGA | A CCC | SATTO | TAA | CGAT | rgcto | GA A | ATGGA | ATGCTG | 120 |
| | | | | | | | | | | | | | | | | |

WO 99/63055 PCT/US99/12121

| TTCAGAGACG ACTCATGTTT GAAGACGAAT GCATTCTCGT TGATGAAAAT AATCGTG | TGG 180 |
|--|--|
| TGGGACATGA CACTAAGTAT AACTGTCATC TGATGGAAAA GATTGAAGCT GAGAATT | TAC 240 |
| TTCACAGAGC TTTCAGTGTG TTTTTATTCA ACTCCAAGTA TGAGTTGCTT CTCCAGC | AAC 300 |
| GGTCAAAAAC AAAGGTTACT TTCCCACTTG TGTGGACAAA CACTTGTTGC AGCCATC | CTC 360 |
| TTTACCGTGA ATCCGAGCTT ATTGAAGAGA ATGTGCTTGG TGTAAGAAAT GCCGCACA | AAA 420 |
| GGAAGCTTTT CGATGAGCTC GGTATTGTAG CAGAAGATGT ACCAGTCGAT GAGTTCA | CTC 480 |
| CCTTGGGACG CATGCTTTAC AAGGCACCTT CTGATGGGAA ATGGGGAGAG CACGAAG | TTG 540 |
| ACTATCTACT CTTCATCGTG CGGGATGTGA AGCTTCAACC AAACCCAGAT GAAGTGG | CTG 600 |
| AGATCAAGTA CGTGAGCAGG GAAGAGCTTA AGGAGCTGGT GAAGAAAGCA GATGCTG | GCG . 660 |
| ATGAAGCTGT GAAACTATCT CCATGGTTCA GATTGGTGGT GGATAATTTC TTGATGA | AGT 720 |
| GGTGGGATCA TGTTGAGAAA GGAACTATCA CTGAAGCTGC AGACATGAAA ACCATTC | ACA 780 |
| AGCTCTGAAC TTTCCATAAG TTTTGGATCT TCCCCTTCCC ATAATAAAAT TAAGAGA | TGA 840 |
| GACTTTTATT GATTACAGAC AAAACTGGCA ACAAAATCTA TTCCTAGGAT TTTTTTT | TGC 900 |
| TTTTTATTTA CTTTTGATTC ATCTCTAGTT TAGTTTTCAT CTTAAAAAAA AAAA | 954 |
| | |
| <210> 10 <211> 996 <212> DNA <213> Arabidopsis thaliana | |
| <211> 996 <212> DNA | CTC 60 |
| <211> 996 <212> DNA <213> Arabidopsis thaliana <400> 10 | |
| <211> 996 <212> DNA <213> Arabidopsis thaliana <400> 10 CACCAATGTC TGTTTCTTCT TTATTTAATC TCCCATTGAT TCGCCTCAGA TCTCTCG | ÇGA 120 |
| <pre><211> 996 <212> DNA <213> Arabidopsis thaliana <400> 10 CACCAATGTC TGTTTCTTCT TTATTTAATC TCCCATTGAT TCGCCTCAGA TCTCTCG TTTCGTCTTC TTTTCTTCT TTCCGATTTG CCCATCGTCC TCTGTCATCG ATTTCAC</pre> | CGA 120 CTG 180 |
| <pre><211> 996 <212> DNA <213> Arabidopsis thaliana <400> 10 CACCAATGTC TGTTTCTTCT TTATTTAATC TCCCATTGAT TCGCCTCAGA TCTCTCG TTTCGTCTTC TTTTCTTCT TTCCGATTTG CCCATCGTCC TCTGTCATCG ATTTCAC GAAAGTTACC GAATTTTCGT GCTTTCTCTG GTACCGCTAT GACAGATACT AAAGATG</pre> | CGA 120 CTG 180 AAA 240 |
| <pre><211> 996 <212> DNA <213> Arabidopsis thaliana <400> 10 CACCAATGTC TGTTTCTTCT TTATTTAATC TCCCATTGAT TCGCCTCAGA TCTCTCG TTTCGTCTTC TTTTCTTCT TTCCGATTTG CCCATCGTCC TCTGTCATCG ATTTCAC GAAAGTTACC GAATTTTCGT GCTTTCTCTG GTACCGCTAT GACAGATACT AAAGATG GTATGGATGC TGTTCAGAGA CGTCTCATGT TTGAGGATGA ATGCATTCTT GTTGATG</pre> | CGA 120 CTG 180 AAA 240 AAG 300 |
| <pre><211> 996 <212> DNA <213> Arabidopsis thaliana <400> 10 CACCAATGTC TGTTTCTTCT TTATTTAATC TCCCATTGAT TCGCCTCAGA TCTCTCG TTTCGTCTTC TTTTCTTCT TTCCGATTTG CCCATCGTCC TCTGTCATCG ATTTCAC GAAAGTTACC GAATTTTCGT GCTTTCTCTG GTACCGCTAT GACAGATACT AAAGATG GTATGGATGC TGTTCAGAGA CGTCTCATGT TTGAGGATGA ATGCATTCTT GTTGATG CTGATCGTGT TGTGGGGCAT GTCAGCAAGT ATAATTGTCA TCTGATGGAA AATATTG</pre> | CGA 120 CTG 180 AAA 240 AAG 300 TGC 360 |
| <pre><211> 996 <212> DNA <213> Arabidopsis thaliana <400> 10 CACCAATGTC TGTTTCTTCT TTATTTAATC TCCCATTGAT TCGCCTCAGA TCTCTCG TTTCGTCTTC TTTTCTTCT TTCCGATTTG CCCATCGTCC TCTGTCATCG ATTTCAC GAAAGTTACC GAATTTTCGT GCTTTCTCTG GTACCGCTAT GACAGATACT AAAGATG GTATGGATGC TGTTCAGAGA CGTCTCATGT TTGAGGATGA ATGCATTCTT GTTGATG CTGATCGTGT TGTGGGGCAT GTCAGCAAGT ATAATTGTCA TCTGATGGAA AATATTG CCAAGAATTT GCTGCACAGG GCTTTTAGTG TATTTTTATT CAACTCGAAG TATGAGT</pre> | CGA 120 CTG 180 AAA 240 AAG 300 TGC 360 GTT 420 |
| <pre><211> 996 <212> DNA <213> Arabidopsis thaliana <400> 10 CACCAATGTC TGTTTCTTCT TTATTTAATC TCCCATTGAT TCGCCTCAGA TCTCTCG TTTCGTCTTC TTTTTCTTCT TTCCGATTTG CCCATCGTCC TCTGTCATCG ATTTCACC GAAAGTTACC GAATTTTCGT GCTTTCTCTG GTACCGCTAT GACAGATACT AAAGATG GTATGGATGC TGTTCAGAGA CGTCTCATGT TTGAGGATGA ATGCATTCTT GTTGATG CTGATCGTGT TGTGGGGCAT GTCAGCAAGT ATAATTGTCA TCTGATGGAA AATATTG CCAAGAATTT GCTGCACAGG GCTTTTAGTG TATTTTTATT CAACTCGAAG TATGAGT TTCTCCAGCA AAGGTCAAAC ACAAAGGTTA CGTTCCCTCT AGTGTGGACT AACACTT</pre> | CGA 120 CTG 180 AAA 240 AAG 300 TGC 360 GTT 420 GGA 480 |
| <pre><211> 996 <212> DNA <213> Arabidopsis thaliana </pre> <pre><400> 10 CACCAATGTC TGTTTCTTCT TTATTTAATC TCCCATTGAT TCGCCTCAGA TCTCTCG TTTCGTCTTC TTTTTCTTCT TTCCGATTTG CCCATCGTCC TCTGTCATCG ATTTCAC GAAAGTTACC GAATTTTCGT GCTTTCTCTG GTACCGCTAT GACAGATACT AAAGATG GTATGGATGC TGTTCAGAGA CGTCTCATGT TTGAGGATGA ATGCATTCTT GTTGATG CTGATCGTGT TGTGGGGCAT GTCAGCAAGT ATAATTGTCA TCTGATGGAA AATATTG CCAAGAATTT GCTGCACAGG GCTTTTAGTG TATTTTTATT CAACTCGAAG TATGAGT TTCTCCAGCA AAGGTCAAAC ACAAAGGTTA CGTTCCCTCT AGTGTGGACT AACACTT GCAGCCATCC TCTTTACCGT GAATCAGAGC TTATCCAGGA CAATGCACTA GGTGTGA</pre> | CGA 120 CTG 180 AAA 240 AAG 300 TGC 360 GTT 420 GGA 480 TCG 540 |
| <pre><211> 996 <212> DNA <213> Arabidopsis thaliana <400> 10 CACCAATGTC TGTTTCTTCT TTATTTAATC TCCCATTGAT TCGCCTCAGA TCTCTCG TTTCGTCTTC TTTTTCTTCT TTCCGATTTG CCCATCGTCC TCTGTCATCG ATTTCAC GAAAGTTACC GAATTTCGT GCTTCTCTG GTACCGCTAT GACAGATACT AAAGATG GTATGGATGC TGTTCAGAGA CGTCTCATGT TTGAGGATGA ATGCATTCTT GTTGATG CTGATCGTGT TGTGGGGCAT GTCAGCAAGT ATAATTGTCA TCTGATGGAA AATATTG CCAAGAATTT GCTGCACAGG GCTTTTAGTG TATTTTTATT CAACTCGAAG TATGAGT TTCTCCAGCA AAGGTCAAAC ACAAAGGTTA CGTTCCCTCT AGTGTGGACT AACACTT GCAGCCATCC TCTTTACCGT GAATCAGAGC TTATCCAGGA CAATGCACTA GGTGTGA ATGCTGCACA AAGAAAGCTT CTCGATGAGC TTGGTATTGT AGCTGAAGAT GTACCAGG</pre> | CGA 120 CTG 180 AAA 240 AAG 300 TGC 360 GTT 420 AGGA 480 GTCG 540 GGAG 600 |

780

CAGATGCAGG TGAGGAAGGT TTGAAACTGT CACCATGGTT CAGATTGGTG GTGGACAATT

| WO 99/63055 | |
|-------------|--|



| | | | | | • | |
|---|-------------|------------|------------|------------|------------|------|
| TCTTGATGAA | GTGGTGGGAT | CATGTTGAGA | AAGGAACTTT | GGTTGAAGCT | ATAGACATGA | 840 |
| AAACCATCCA | CAAACTCTGA | ACATCTTTTT | TTAAAGTTTT | ТАААТСААТС | AACTTTCTCT | 900 |
| TCATCATTTT | TATCTTTTCG | ATGATAATAA | TTTGGGATAT | GTGAGACACT | TACAAAACTT | 960 |
| CCAAGCACCT | CAGGCAATAA | TAAAGTTTGC | GGCCGC | | | 996 |
| <210> 11 <211> 1165 <212> DNA <213> Haem | atococcus p | luvialis | | | | |
| <400> 11 CTCGGTAGCT | GGCCACAATC | GCTATTTGGA | ACCTGGCCCG | GCGGCAGTCC | GATGCCGCGA | 60 |
| TGCTTCGTTC | GTTGCTCAGA | GGCCTCACGC | ATATCCCCCG | CGTGAACTCC | GCCCAGCAGC | 120 |
| CCAGCTGTGC | ACACGCGCGA | CTCCAGTTTA | AGCTCAGGAG | CATGCAGATG | ACGCTCATGC | 180 |
| AGCCCAGCAT | CTCAGCCAAT | CTGTCGCGCG | CCGAGGACCG | CACAGACCAC | ATGAGGGGTG | 240 |
| CAAGCACCTG | GGCAGGCGGG | CAGTCGCAGG | ATGAGCTGAT | GCTGAAGGAC | GAGTGCATCT | 300 |
| TGGTGGATGT | TGAGGACAAC | ATCACAGGCC | ATGCCAGCAA | GCTGGAGTGT | CACAAGTTCC | 360 |
| TACCACATCA | GCCTGCAGGC | CTGCTGCACC | GGGCCTTCTC | TGTGTTCCTG | TTTGACGATC | 420 |
| AGGGGCGACT | GCTGCTGCAA | CAGCGTGCAC | GCTCAAAAAT | CACCTTCCCA | AGTGTGTGGA | 480 |
| CGAACACCTG | CTGCAGCCAC | CCTTTACATG | GGCAGACCCC | AGATGAGGTG | GACCAACTAA | 540 |
| GCCAGGTGGC | CGACGGAACA | GTACCTGGCG | CAAAGGCTGC | TGCCATCCGC | AAGTTGGAGC | 600 |
| ACGAGCTGGG | GATACCAGCG | CACCAGCTGC | CGGCAAGCGC | GTTTCGCTTC | CTCACGCGTT | 660 |
| TGCACTACTG | TGCCGCGGAC | GTGCAGCCAG | CTGCGACACA | ATCAGCGCTC | TGGGGCGAGC | 720 |
| ACGAAATGGA | CTACATCTTG | TTCATCCGGG | CCAACGTCAC | CTTGGCGCCC | AACCCTGACG | 780 |
| AGGTGGACGA | AGTCAGGTAC | GTGACGCAAG | AGGAGCTGCG | GCAGATGATG | CAGCCGGACA | 840 |
| ACGGGCTGCA | ATGGTCGCCG | TGGTTTCGCA | TCATCGCCGC | GCGCTTCCTT | GAGCGTTGGT | 900 |
| GGGCTGACCT | GGACGCGGCC | CTAAACACTG | ACAAACACGA | GGATTGGGGA | ACGGTGCATC | 960 |
| ACATCAACGA | AGCGTGAAAG | CAGAAGCTGC | AGGATGTGAA | GACACGTCAT | GGGGTGGAAT | 1020 |
| TGCGTACTTG | GCAGCTTCGT | ATCTCCTTTT | TCTGAGACTG | AACCTGCAGT | CAGGTCCCAC | 1080 |
| AAGGTCAGGT | AAAATGGCTC | GATAAAATGT | ACCGTCACTT | TTTGTCGCGT | ATACTGAACT | 1140 |
| CCAAGAGGTC | AAAAAAAA | AAAAA | | | | 1165 |

<210> 12 <211> 1135 <212> DNA

<213> Haematococcus pluvialis

| WO 99/63055 | |
|-------------|--|

| <400> 12 CTCGGTAGCT | GGCCACAATC | GCTATTTGGA | ACCTGGCCCG | GCGGCAGTCC | GATGCCGCGA | 60 |
|---|------------|-------------|------------|------------|------------|------|
| TGCTTCGTTC | GTTGCTCAGA | GGCCTCACGC | ATATCCCGCG | CGTGAACTCC | GCCCAGCAGC | 120 |
| CCAGCTGTGC | ACACGCGCGA | CTCCAGTTTA. | AGCTCAGGAG | CATGCAGCTG | CTTTCCGAGG | 180 |
| ACCGCACAGA | CCACATGAGG | GGTGCAAGCA | CCTGGGCAGG | CGGGCAGTCG | CAGGATGAGC | 240 |
| TGATGCTGAA | GGACGAGTGC | ATCTTGGTAG | ATGTTGAGGA | CAACATCACA | GGCCATGCCA | 300 |
| GCAAGCTGGA | GTGTCACAAG | TTCCTACCAC | ATCAGCCTGC | AGGCCTGCTG | CACCGGGCCT | 360 |
| TCTCTGTGTT | CCTGTTTGAC | GATCAGGGGC | GACTGCTGCT | GCAACAGCGT | GCACGCTCAA | 420 |
| AAATCACCTT | CCCAAGTGTG | TGGACGAACA | CCTGCTGCAG | CCACCCTTTA | CATGGGCAGA | 480 |
| CCCCAGATGA | GGTGGACCAA | CTAAGCCAGG | TGGCCGACGG | AACAGTACCT | GGCGCAAAGG | 540 |
| CTGCTGCCAT | CCGCAAGTTG | GAGCACGAGC | TGGGGATACC | AGCGCACCAG | CTGCCGGCAA | 600 |
| GCGCGTTTCG | CTTCCTCACG | CGTTTGCACT | ACTGTGCCGC | GGACGTGCAG | CCAGCTGCGA | 660 |
| CACAATCAGC | GCTCTGGGGC | GAGCACGAAA | TGGACTACAT | CTTGTTCATC | CGGGCCAACG | 720 |
| TCACCTTGGC | GCCCAACCCT | GACGAGGTGG | ACGAAGTCAG | GTACGTGACG | CAAGAGGAGC | 780 |
| TGCGGCAGAT | GATGCAGCCG | GACAACGGGC | TTCAATGGTC | GCCGTGGTTT | CGCATCATCG | 840 |
| CCGCGCGCTT | CCTTGAGCGT | TGGTGGGCTG | ACCTGGACGC | GGCCCTAAAC | ACTGACAAAC | 900 |
| ACGAGGATTG | GGGAACGGTG | CATCACATCA | ACGAAGCGTG | AAGGCAGAAG | CTGCAGGATG | 960 |
| TGAAGACACG | TCATGGGGTG | GAATTGCGTA | CTTGGCAGCT | TCGTATCTCC | TTTTTCTGAG | 1020 |
| ACTGAACCTG | CAGAGCTAGA | GTCAATGGTG | CATCATATTC | ATCGTCTCTC | TTTTGTTTTA | 1080 |
| GACTAATCTG | TAGCTAGAGT | CACTGATGAA | TCCTTTACAA | CTTTCAAAAA | AAAAA | 1135 |
| <210> 13 <211> 960 <212> DNA <213> Taget | tes erecta | | · | | | |
| <400> 13 | | | | | | |
| | CTCAAATCTC | • | | • | | 60 |
| | TCAGCGACGT | | | | | 120 |
| | GGGACATGAT | | | | | 180 |
| | GCACAGAGCA | | | | | 240 |
| | GTCTGCAACC | | | | | 300 |
| * | CTACAGAGAA | | | | | 360 |
| | NNNNNNNNN | | | | - | 420 |
| иииииииии | иииииииии | иииииииии | иииииииии | иииииииии | NNNNNNNNN | 480 |



| | | | | | - | |
|------------|------------|------------|------------|------------|------------|-----|
| иииииииии | иииииииии | иииииииии | иииииииии | иииииииии | иииииииии | 540 |
| имимимими | имимимими | имимимими | имимимими | иииииииии | имимимими | 600 |
| ииииииииии | иииииииии | ииииииииии | имимимими | ииииииииии | ииииииииии | 660 |
| ииииииииии | иииииииии | TCATGTGCAA | AAGGGTACAC | TCACTGAATG | CAATTTGATA | 720 |
| TGAAAACCAT | ACACAAGCTG | ATATAGAAAC | ACACCCTCAA | CCGAAAAGCA | AGCCTAATAA | 780 |
| TTCGGGTTGG | GTCGGGTCTA | CCATCAATTG | TTTTTTTTTT | TTAACAACTT | TTAATCTCTA | 840 |
| TTTGAGCATG | TTGATTCTTG | TCTTTTGTGT | GTAAGATTTT | GGGTTTCGTT | TCAGTTGTAA | 900 |
| TAATGAACCA | TTGATGGTTT | GCAATTTCAA | GTTCCTATCG | ACATGTAGTG | АТСТАААААА | 960 |
| | | | | | | |

<210> 14

<211> 305

<212> PRT

<213> Haematococcus pluvialis

<400> 14

Met Leu Arg Ser Leu Leu Arg Gly Leu Thr His Ile Pro Arg Val Asn
1 5 10 15

Ser Ala Gln Gln Pro Ser Cys Ala His Ala Arg Leu Gln Phe Lys Leu 20 25 30

Arg Ser Met Gln Met Thr Leu Met Gln Pro Ser Ile Ser Ala Asn Leu 35 40 45

Ser Arg Ala Glu Asp Arg Thr Asp His Met Arg Gly Ala Ser Thr Trp 50 55 60

Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys Asp Glu Cys Ile 65 70 75 80

Leu Val Asp Val Glu Asp Asn Ile Thr Gly His Ala Ser Lys Leu Glu 85 90 95

Cys His Lys Phe Leu Pro His Gln Pro Ala Gly Leu Leu His Arg Ala 100 105 110

Phe Ser Val Phe Leu Phe Asp Asp Gln Gly Arg Leu Leu Gln Gln 115 120 125

Arg Ala Arg Ser Lys Ile Thr Phe Pro Ser Val Trp Thr Asn Thr Cys 130 135 140

Cys Ser His Pro Leu His Gly Gln Thr Pro Asp Glu Val Asp Gln Leu 145 150 155 160

Ser Gln Val Ala Asp Gly Thr Val Pro Gly Ala Lys Ala Ala Ile 165 170 175

Arg Lys Leu Glu His Glu Leu Gly Ile Pro Ala His Gln Leu Pro Ala 180 185 190

Ser Ala Phe Arg Phe Leu Thr Arg Leu His Tyr Cys Ala Ala Asp Val 195 200 205



Gln Pro Ala Ala Thr Gln Ser Ala Leu Trp Gly Glu His Glu Met Asp 210 215 220

Tyr Ile Leu Phe Ile Arg Ala Asn Val Thr Leu Ala Pro Asn Pro Asp 225 230 235 240

Glu Val Asp Glu Val Arg Tyr Val Thr Gln Glu Glu Leu Arg Gln Met 245 250 255

Met Gln Pro Asp Asn Gly Leu Gln Trp Ser Pro Trp Phe Arg Ile Ile 260 265 270

Ala Ala Arg Phe Leu Glu Arg Trp Trp Ala Asp Leu Asp Ala Ala Leu 275 280 285

Asn Thr Asp Lys His Glu Asp Trp Gly Thr Val His His Ile Asn Glu 290 295 300

Ala 305

<210> 15

<211> 293

<212> PRT

<213> Haematococcus pluvialis

<400> 15

Met Leu Arg Ser Leu Leu Arg Gly Leu Thr His Ile Pro Arg Val Asn 1 5 10 15

Ser Ala Gln Gln Pro Ser Cys Ala His Ala Arg Leu Gln Phe Lys Leu 20 25 30

Arg Ser Met Gln Leu Leu Ser Glu Asp Arg Thr Asp His Met Arg Gly
35 40 45

Ala Ser Thr Trp Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys
50 55 60

Asp Glu Cys Ile Leu Val Asp Val Glu Asp Asn Ile Thr Gly His Ala 65 70 75 80

Ser Lys Leu Glu Cys His Lys Phe Leu Pro His Gln Pro Ala Gly Leu 85 90 95

Leu His Arg Ala Phe Ser Val Phe Leu Phe Asp Asp Gln Gly Arg Leu 100 105 110

Leu Leu Gln Gln Arg Ala Arg Ser Lys Ile Thr Phe Pro Ser Val Trp

Thr Asn Thr Cys Cys Ser His Pro Leu His Gly Gln Thr Pro Asp Glu 130 135 140

Val Asp Gln Leu Ser Gln Val Ala Asp Gly Thr Val Pro Gly Ala Lys
145 150 155 160

Ala Ala Ala Ile Arg Lys Leu Glu His Glu Leu Gly Ile Pro Ala His 165 170 175

Gln Leu Pro Ala Ser Ala Phe Arg Phe Leu Thr Arg Leu His Tyr Cys

180 185 190

Ala Ala Asp Val Gln Pro Ala Ala Thr Gln Ser Ala Leu Trp Gly Glu 195 200 205

His Glu Met Asp Tyr Ile Leu Phe Ile Arg Ala Asn Val Thr Leu Ala 210 215 220

Pro Asn Pro Asp Glu Val Asp Glu Val Arg Tyr Val Thr Gln Glu Glu 225 230 235 240

Leu Arg Gln Met Met Gln Pro Asp Asn Gly Leu Gln Trp Ser Pro Trp 245 250 255

Phe Arg Ile Ile Ala Ala Arg Phe Leu Glu Arg Trp Trp Ala Asp Leu 260 265 270

Asp Ala Ala Leu Asn Thr Asp Lys His Glu Asp Trp Gly Thr Val His 275 280 285

His Ile Asn Glu Ala 290

<210> 16

<211> 284

<212> PRT

<213> Arabidopsis thaliana

<400> 16

Met Ser Val Ser Ser Leu Phe Asn Leu Pro Leu Ile Arg Leu Arg Ser 1 5 10 15

Leu Ala Leu Ser Ser Ser Phe Ser Ser Phe Arg Phe Ala His Arg Pro
20 25 30

Leu Ser Ser Ile Ser Pro Arg Lys Leu Pro Asn Phe Arg Ala Phe Ser 35 40 45

Gly Thr Ala Met Thr Asp Thr Lys Asp Ala Gly Met Asp Ala Val Gln 50 60

Arg Arg Leu Met Phe Glu Asp Glu Cys Ile Leu Val Asp Glu Thr Asp 65 70 75 80

Arg Val Val Gly His Val Ser Lys Tyr Asn Cys His Leu Met Glu Asn 85 90 95

Ile Glu Ala Lys Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe 100 105 110

Asn Ser Lys Tyr Glu Leu Leu Gln Gln Arg Ser Asn Thr Lys Val 115 120 125

Thr Phe Pro Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr 130 135 140

Arg Glu Ser Glu Leu Ile Gln Asp Asn Ala Leu Gly Val Arg Asn Ala 145 150 . 155 160

Ala Gln Arg Lys Leu Leu Asp Glu Leu Gly Ile Val Ala Glu Asp Val 165 170 175



Pro Val Asp Glu Phe Thr Pro Leu Gly Arg Met Leu Tyr Lys Ala Pro 180 185 190

Ser Asp Gly Lys Trp Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Ile 195 200 205

Val Arg Asp Val Lys Val Gln Pro Asn Pro Asp Glu Val Ala Glu Ile 210 215 220

Lys Tyr Val Ser Arg Glu Glu Leu Lys Glu Leu Val Lys Lys Ala Asp 225 230 235 240

Ala Gly Glu Glu Leu Lys Leu Ser Pro Trp Phe Arg Leu Val Val 245 250 255

Asp Asn Phe Leu Met Lys Trp Trp Asp His Val Glu Lys Gly Thr Leu 260 265 270

Val Glu Ala Ile Asp Met Lys Thr Ile His Lys Leu 275 280

<210> 17

<211> 287

<212> PRT

<213> Clarkia breweri

<400> 17

Met Ser Ser Ser Met Leu Asn Phe Thr Ala Ser Arg Ile Val Ser Leu
1 5 10 15

Pro Leu Eur Ser Pro Pro Ser Arg Val His Leu Pro Leu Cys Phe 20 25 30

Phe Ser Pro Ile Ser Leu Thr Gln Arg Phe Ser Ala Lys Leu Thr Phe 35 40 45

Ser Ser Gln Ala Thr Thr Met Gly Glu Val Val Asp Ala Gly Met Asp 50 55 60

Ala Val Gln Arg Arg Leu Met Phe Glu Asp Glu Cys Ile Leu Val Asp 65 70 75 80

Glu Asn Asp Lys Val Val Gly His Glu Ser Lys Tyr Asn Cys His Leu 85 90 95

Met Glu Lys Ile Glu Ser Glu Asn Leu Leu His Arg Ala Phe Ser Val 100 105 110

Phe Leu Phe Asn Ser Lys Tyr Glu Leu Leu Gln Gln Arg Ser Ala 115 120 125

Thr Lys Val Thr Phe Pro Leu Val Trp Thr Asn Thr Cys Cys Ser His 130 135 140

Pro Leu Tyr Arg Glu Ser Glu Leu Ile Asp Glu Asn Cys Leu Gly Val 145 150 155 160

Arg Asn Ala Ala Gln Arg Lys Leu Leu Asp Glu Leu Gly Ile Pro Ala 165 170 175

Glu Asp Leu Pro Val Asp Gln Phe Ile Pro Leu Ser Arg Ile Leu Tyr

190

180 185

Lys Ala Pro Ser Asp Gly Lys Trp Gly Glu His Glu Leu Asp Tyr Leu 195 200 205

Leu Phe Ile Ile Arg Asp Val Asn Leu Asp Pro Asn Pro Asp Glu Val 210 215 220

Ala Glu Val Lys Tyr Met Asn Arg Asp Asp Leu Lys Glu Leu Leu Arg 225 230 235 240

Lys Ala Asp Ala Glu Glu Glu Gly Val Lys Leu Ser Pro Trp Phe Arg 245 250 255

Leu Val Val Asp Asn Phe Leu Phe Lys Trp Trp Asp His Val Glu Lys 260 265 270

Gly Ser Leu Lys Asp Ala Ala Asp Met Lys Thr Ile His Lys Leu 275 280 285

<210> 18

<211> 261

<212> PRT

<213> Arabidopsis thaliana

<400> 18

Thr Gly Pro Pro Pro Arg Phe Phe Pro Ile Arg Ser Pro Val Pro Arg
1 5 10 15

Thr Gln Leu Phe Val Arg Ala Phe Ser Ala Val Thr Met Thr Asp Ser 20 25 30

Asn Asp Ala Gly Met Asp Ala Val Gln Arg Arg Leu Met Phe Glu Asp 35 40 45

Glu Cys Ile Leu Val Asp Glu Asn Asn Arg Val Val Gly His Asp Thr 50 55 60

Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala Glu Asn Leu Leu 65 70 75 80

His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu Leu Leu 85 90 95

Leu Gln Gln Arg Ser Lys Thr Lys Val Thr Phe Pro Leu Val Trp Thr 100 105 110

Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser Glu Leu Ile Glu 115 120 125

Glu Asn Val Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu Phe Asp 130 135 140

Glu Leu Gly Ile Val Ala Glu Asp Val Pro Val Asp Glu Phe Thr Pro 145 150 155 160

Leu Gly Arg Met Leu Tyr Lys Ala Pro Ser Asp Gly Lys Trp Gly Glu
165

His Glu Val Asp Tyr Leu Leu Phe Ile Val Arg Asp Val Lys Leu Gln 180 185 190





Pro Asn Pro Asp Glu Val Ala Glu Ile Lys Tyr Val Ser Arg Glu Glu 195 200 205

Leu Lys Glu Leu Val Lys Lys Ala Asp Ala Gly Asp Glu Ala Val Lys 210 220

Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe Leu Met Lys Trp 225 230 235 240

Trp Asp His Val Glu Lys Gly Thr Ile Thr Glu Ala Ala Asp Met Lys 245 250 255

Thr Ile His Lys Leu 260

<210> 19

<211> 288

<212> PRT

<213> Saccharomyces cerevisiae

<400> 19

Met Thr Ala Asp Asn Asn Ser Met Pro His Gly Ala Val Ser Ser Tyr 1 5 10 15

Ala Lys Leu Val Gln Asn Gln Thr Pro Glu Asp Ile Leu Glu Glu Phe 20 25 30

Pro Glu Ile Ile Pro Leu Gln Gln Arg Pro Asn Thr Arg Ser Ser Glu 35 40 45

Thr Ser Asn Asp Glu Ser Gly Glu Thr Cys Phe Ser Gly His Asp Glu 50 55 60

Glu Gln Ile Lys Leu Met Asn Glu Asn Cys Ile Val Leu Asp Trp Asp 65 70 75 80

Asp Asn Ala Ile Gly Ala Gly Thr Lys Lys Val Cys His Leu Met Glu 85 90 95

Asn Ile Glu Lys Gly Leu Leu His Arg Ala Phe Ser Val Phe Ile Phe 100 105 110

Asn Glu Gln Gly Glu Leu Leu Gln Gln Arg Ala Thr Glu Lys Ile 115 120 125

Thr Phe Pro Asp Leu Trp Thr Asn Thr Cys Cys Ser His Pro Leu Cys 130 135 140

Ile Asp Asp Glu Leu Gly Leu Lys Gly Lys Leu Asp Asp Lys Ile Lys 145 150 155 160

Gly Ala Ile Thr Ala Ala Val Arg Lys Leu Asp His Glu Leu Gly Ile 165 170 175

Pro Glu Asp Glu Thr Lys Thr Arg Gly Lys Phe His Phe Leu Asn Arg 180 185 190

Ile His Tyr Met Ala Pro Ser Asn Glu Pro Trp Gly Glu His Glu Ile 195 200 205

Asp Tyr Ile Leu Phe Tyr Lys Ile Asn Ala Lys Glu Asn Leu Thr Val

210 215 220

Asn Pro Asn Val Asn Glu Val Arg Asp Phe Lys Trp Val Ser Pro Asn 225 235 240

Asp Leu Lys Thr Met Phe Ala Asp Pro Ser Tyr Lys Phe Thr Pro Trp 245 250 255

Phe Lys Ile Ile Cys Glu Asn Tyr Leu Phe Asn Trp Trp Glu Gln Leu 260 265 270

Asp Asp Leu Ser Glu Val Glu Asn Asp Arg Gln Ile His Arg Met Leu 275 280 285

<210> 20

<211> 456

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Consensus
 sequence of four plant B-cyclases

<400> 20

Met Asp Thr Leu Leu Lys Thr Pro Asn Leu Glu Phe Leu Pro His Gly
1 5 10 15

Phe Val Lys Ser Phe Ser Lys Phe Gly Lys Cys Glu Gly Val Cys Val 20 25 30

Lys Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr Lys Lys Glu Asn 35 40 45

Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys Gly Val Val Asp
50 55 60

Leu Ala Val Val Gly Gly Pro Ala Gly Leu Ala Val Ala Gln Gln 65 70 75 80'

Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile Asp Pro Pro Lys Leu 85 90 95

Ile Trp Pro Asn Asn Tyr Gly Val Trp Val Asp Glu Phe Glu Ala Met 100 105 110

Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp Ser Gly Ala Val Tyr Ile 115 120 125

Asp Asp Thr Lys Asp Leu Arg Pro Tyr Gly Arg Val Asn Arg Lys Gln 130 135 140

Leu Lys Ser Lys Met Met Gln Lys Cys Ile Asn Gly Val Lys Phe His 145 150 155 160

Gln Ala Lys Val Ile Lys Val Ile His Glu Glu Lys Ser Met Leu Ile 165 170 175

Cys Asn Asp Gly Thr Ile Gln Ala Thr Val Val Leu Asp Ala Thr Gly 180 185 190

Phe Ser Arg Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln

, ·

195

200 205

Val Ala Tyr Gly Ile Leu Ala Glu Val Glu Glu His Pro Phe Asp Lys 210 215 220

Met Val Phe Met Asp Trp Arg Asp Ser His Leu Asn Asn Glu Leu Lys 235 240

Glu Arg Asn Ser Ile Pro Thr Phe Leu Tyr Ala Met Pro Phe Ser Ser 245 250 255

Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg Pro Gly Leu 260 265 270

Arg Met Asp Asp Ile Gln Glu Arg Met Val Ala Arg Leu His Leu Gly 275 280 285

Ile Lys Val Lys Ser Ile Glu Glu Asp Glu His Cys Val Ile Pro Met 290 295 300

Gly Gly Pro Leu Pro Val Leu Pro Gln Arg Val Val Gly Ile Gly Gly 305 310 315 320

Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met Val Ala Arg Thr 325 330 335

Leu Ala Ala Pro Val Val Ala As
n Ala Ile Ile Tyr Leu Gly Ser 340 345 350

Glu Ser Ser Gly Glu Leu Ser Ala Glu Val Trp Lys Asp Leu Trp Pro 355 360 365

Ile Glu Arg Arg Arg Gln Arg Glu Phe Phe Cys Phe Gly Met Asp Ile 370 375 380

Leu Leu Lys Leu Asp Leu Pro Ala Thr Arg Arg Phe Phe Asp Ala Phe 385 390 395 400

Phe Asp Leu Glu Pro Arg Tyr Trp His Gly Phe Leu Ser Ser Arg Leu 405 410 415

Phe Leu Pro Glu Leu Ile Val Phe Gly Leu Ser Leu Phe Ser His Ala 420 425 430

Ser Asn Thr Ser Arg Glu Ile Met Thr Lys Gly Thr Pro Leu Val Met 435 440 445

Ile Asn Asn Leu Leu Gln Asp Glu 450 455

<210> 21

<211> 524

<212> PRT

<213> Arabidopsis thaliana

<400> 21

Met Glu Cys Val Gly Ala Arg Asn Phe Ala Ala Met Ala Val Ser Thr 1 5 10 15

Phe Pro Ser Trp Ser Cys Arg Arg Lys Phe Pro Val Val Lys Arg Tyr 20 25 30

WO 99/63055

Ser Tyr Arg Asn Ile Arg Phe Gly Leu Cys Ser Val Arg Ala Ser Gly 40 Gly Gly Ser Ser Gly Ser Glu Ser Cys Val Ala Val Arg Glu Asp Phe Ala Asp Glu Glu Asp Phe Val Lys Ala Gly Gly Ser Glu Ile Leu Phe Val Gln Met Gln Gln Asn Lys Asp Met Asp Glu Gln Ser Lys Leu Val Asp Lys Leu Pro Pro Ile Ser Ile Gly Asp Gly Ala Leu Asp His Val Val Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala Ala Glu Ser Ala 120 Lys Leu Gly Leu Lys Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Glu Asp Glu Phe Asn Asp Leu Gly Leu Gln Lys Cys Ile Glu His Val Trp Arg Glu Thr Ile Val Tyr Leu Asp Asp 170 Asp Lys Pro Ile Thr Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg Arg Leu Leu His Glu Glu Leu Leu Arg Arg Cys Val Glu Ser Gly Val Ser Tyr Leu Ser Ser Lys Val Asp Ser Ile Thr Glu Ala Ser Asp Gly Leu Arg Leu Val Ala Cys Asp Asp Asn Asn Val Ile Pro Cys Arg Leu Ala 235 Thr Val Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Gln Tyr Glu Val Gly Gly Pro Arg Val Cys Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Glu Asn Ser Pro Tyr Asp Pro Asp Gln Met Val Phe Met Asp Tyr Arg Asp Tyr Thr Asn Glu Lys Val Arg Ser Leu Glu Ala Glu Tyr Pro Thr Phe Leu Tyr Ala Met Pro Met Thr Lys Ser Arg Leu Phe Phe Glu 310 315 Glu Thr Cys Leu Ala Ser Lys Asp Val Met Pro Phe Asp Leu Leu Lys 330 Thr Lys Leu Met Leu Arg Leu Asp Thr Leu Gly Ile Arg Ile Leu Lys 345 Thr Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro



| | | 355 | | | | | 360 | | | | | 505 | | | |
|-----|------------|-----|-----|-----|-----|------------|-----|-----|------|-----|------------|-----|-----|-----|-----|
| Asn | Thr 370 | Glu | Gln | Lys | Asn | Leu 375 | Ala | Phe | Gly. | Ala | Ala 380 | Ala | Ser | Met | Val |

265

His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro 385 390 395 400

Lys Tyr Ala Ser Val Ile Ala Glu Ile Leu Arg Glu Glu Thr Thr Lys 405 410 415

Gln Ile Asn Ser Asn Ile Ser Arg Gln Ala Trp Asp Thr Leu Trp Pro 420 425 430

Pro Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu Phe Gly Leu Ala Leu 435 440 445

Ile Val Gln Phe Asp Thr Glu Gly Ile Arg Ser Phe Phe Arg Thr Phe 450 455 460

Phe Arg Leu Pro Lys Trp Met Trp Gln Gly Phe Leu Gly Ser Thr Leu 465 470 475 480

Thr Ser Gly Asp Leu Val Leu Phe Ala Leu Tyr Met Phe Val Ile Ser 485 490 495

Pro Asn Asn Leu Arg Lys Gly Leu Ile Asn His Leu Ile Ser Asp Pro 500 505 510

Thr Gly Ala Thr Met Ile Lys Thr Tyr Leu Lys Val 515 520

<210> 22

<211> 1898

<212> DNA

<213> Adonis palaestina

<400> 22 AAAGGAGTGT TCTATTAATG TTACTGTCGC ATTCTTGCAA CACTTATATT CAAACTCCAT 60 TTTCTTCTTT TCTCTTCAAA ACAACAAACT AATGTGAGCA GAGTATCTGG CTATGGAACT 120 ACTTGGTGTT CGCAACCTCA TCTCTTCTTG CCCTGTGTGG ACTTTTGGAA CAAGAAACCT 180 TAGTAGTTCA AAACTAGCTT ATAACATACA TCGATATGGT TCTTCTTGTA GAGTAGATTT 240 TCAAGTGAGA GCTGATGGTG GAAGCGGGAG TAGAAGTTCT GTTGCTTATA AAGAGGGTTT 300 TGTGGATGAA GAGGATTTTA TCAAAGCTGG TGGTTCTGAG CTTTTGTTTG TCCAAATGCA 360 GCAAACAAAG TCTATGGAGA AACAGGCCAA GCTCGCCGAT AAGTTGCCAC CAATACCTTT 420 TGGAGAATCC GTGATGGACT TGGTTGTAAT AGGTTGTGGA CCTGCTGGTC TTTCACTGGC 480 TGCAGAAGCT GCTAAGCTAG GGTTGAAAGT TGGCCTTATT GGTCCTGATC TTCCTTTTAC 540 AAATAATTAT GGTGTGGGG AAGACGAGTT CAAAGATCTT GGACTTGAAC GTTGTATCGA 600 GCATGCTTGG AAGGACACCA TCGTATATCT TGATAATGAT GCTCCTGTCC TTATTGGTCG 660 TGCATATGGA CGAGTTAGTC GACATTTGCT ACATGAGGAG TTGCTGAAAA GGTGTGTGGA 720

WO 99/63055 PCT/US99/12121

| | | | | | ÷ '• | |
|------------|------------|------------|------------|------------|------------|------|
| GTCAGGTGTA | TCATATCTTG | ATTCTAAAGT | GGAAAGGATC | ACTGAAGCTG | GTGATGGCCA | 780 |
| TAGCCTTGTA | GTTTGTGAAA | ATGAGATCTT | TATCCCTTGC | AGGCTTGCTA | CTGTTGCATC | 840 |
| TGGAGCAGCT | TCAGGGAAAC | TTTTGGAGTA | TGAAGTAGGT | GGCCCTCGTG | TTTGTGTCCA | 900 |
| AACCGCTTAT | GGGGTGGAGG | TTGAGGTGGA | GAACAATCCA | TACGATCCCA | ACTTAATGGT | 960 |
| ATTCATGGAC | TACAGAGACT | ATATGCAACA | GAAATTACAG | TGCTCGGAAG | AAGAATATCC | 1020 |
| AACATTTCTC | TATGTCATGC | CCATGTCGCC | AACAAGACTT | TTTTTTGAGG | AAACCTGTTT | 1080 |
| GGCCTCAAAA | GATGCCATGC | CATTCGATCT | ACTGAAGAGA | AAACTGATGT | CACGATTGAA | 1140 |
| GACTCTGGGT | ATCCAAGTTA | CAAAAGTTTA | TGAAGAGGAA | TGGTCATATA | TTCCTGTTGG | 1200 |
| TGGTTCTTTA | CCAAACACAG | AGCAAAAGAA | CCTAGCATTT | GGTGCTGCAG | CAAGCATGGT | 1260 |
| GCATCCAGCA | ACAGGCTATT | CGGTTGTACG | GTCACTGTCA | GAAGCTCCAA | AATATGCTTC | 1320 |
| TGTAATTGCA | AAGATTTTGA | AGCAAGATAA | CTCTGCGTAT | GTGGTTTCTG | GACAAAGTAG | 1380 |
| TGCAGTAAAC | ATTTCAATGC | AAGCATGGAG | CAGTCTTTGG | CCAAAGGAGC | GAAAACGTCA | 1440 |
| AAGAGCATTC | TTTCTTTTTG | GATTAGAGCT | TATTGTGCAG | CTAGATATTG | AAGCAACCAG | 1500 |
| AACATTCTTT | AGAACCTTCT | TCCGCTTGCC | AACTTGGATG | TGGTGGGGTT | TCCTTGGGTC | 1560 |
| TTCACTATCA | TCTTTCGATC | TCGTCTTGTT | TTCCATGTAC | ATGTTTGTTT | TGGCGCCAAA | 1620 |
| CAGCATGAGG | ATGTCACTTG | TGAGACATTT | GCTTTCAGAT | CCTTCTGGTG | CAGTTATGGT | 1680 |
| AAGAGCTTAC | CTCGAAAGGT | AGTCTCATCT | ATTATTAAAC | TCTAGTGTTT | CACCAAATAA | 1740 |
| ATGAGGATCC | TTCGAATGTG | TATATGATCA | TCTCTATGTA | TATCCTGTAC | TCTAATCTCA | 1800 |
| TAAAGTAAAT | GCCGGGTTTG | ATATTGTTGT | GTCAAACCGG | CCAATGATAT | AAAGTAAATT | 1860 |
| TATTGATACA | AAAGTAGTTT | TTTTCCTTAA | AAAAAAA | | | 1898 |

<210> 23

<400> 23

Met Glu Leu Leu Gly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp 1 5 10 15

Thr Phe Gly Thr Arg Asn Leu Ser Ser Ser Lys Leu Ala Tyr Asn Ile 20 25 30

His Arg Tyr Gly Ser Ser Cys Arg Val Asp Phe Gln Val Arg Ala Asp 35 40 45

Gly Gly Ser Gly Ser Arg Ser Ser Val Ala Tyr Lys Glu Gly Phe Val 50 55 60

Asp Glu Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu Leu Leu Phe Val 65 70 75 80

<211> 529

<212> PRT

<213> Adonis palaestina



| | | | | | | | | | ~ | | | | | | |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Gln | Met | Gln | Gln | Thr 85 | Lys | Ser | Met | Glu | Lys 90 | Gln | Ala | Lys | Leu | Ala 95 | Asp |
| Lys | Leu | Pro | Pro 100 | Ile | Pro | Phe | Gly | Glu 105 | Ser | «Val | Met | Asp | Leu 110 | Val | Val |
| Ile | Gly | Cys 115 | Gly | Pro | Ala | Gly | Leu 120 | Ser | Leu | Ala | Ala | Glu 125 | Ala | Ala | Lys |
| Leu | Gly 130 | Leu | Lys | Val | Gly | Leu 135 | Ile | Gly | Pro | Asp | Leu 140 | Pro | Phe | Thr | Asn |
| Asn 145 | Tyr | Gly | Val | Trp | Glu 150 | Asp | Glu | Phe | Lys | Asp 155 | Leu | Gly | Leu | Glu | Arg 160 |
| Cys | Ile | Glu | His | Ala 165 | Trp | Lys | Asp | Thr | Ile 170 | Val | Tyr | Leu | Asp | Asn 175 | Asp |
| Ala | Pro | Val | Leu 180 | Ile | Gly | Arg | Ala | Tyr 1 | Gly | Arg | Val | Ser | Arg 190 | His | Leu |
| Leu | His | Glu 195 | Glu | Leu | Leu | Lys | Arg 200 | Cys | Val, | Glu | Ser | Gly 205 | Val | Ser | Tyr |
| Leu | Asp 210 | Ser | Lys | Val | Glu | Arg 215 | Ile | Thr | Glu | Ala | Gly 220 | Asp | Gly | His | Ser |
| Leu 225 | Val | Val | Cys | Glu | Asn 230 | Glu | Ile | Phe | Ile | Pro 235 | Cys | Arg | Leu | Ala | Thr 240 |
| Val | Ala | Ser | Gly | Ala 245 | Ala | Ser | Gly | Lys | Leu 250 | Leu | Glu | Tyr | Glu | Val 255 | Gly |
| Gly | Pro | Arg | Val 260 | Cys | Val | Gln | Thr | Ala 265 | Tyr | Gly | Val | Glu | Val 270 | Glu | Val |
| Glu | Asn | Asn 275 | Pro | Tyr | Asp | Pro | Asn 280 | Leu | Met | Val | Phe | Met 285 | Asp | Tyr | Arg |
| Asp | Tyr 290 | Met | Gln | Gln | Lys | Leu 295 | Gln | Cys | Ser | Glu | Glu 300 | Glu | Tyr | Pro | Thr |
| Phe 305 | Leu | Tyr | Val | Met | Pro 310 | Met | Ser | Pro | Thr | Arg 315 | Leu | Phe | Phe | Glu | Glu 320 |
| Thr | Cys | Leu | Ala | Ser 325 | Lys | Asp | Ala | Met | Pro 330 | Phe | Asp | Leu | Leu | Lys 335 | Arg |
| Lys | Leu | Met | Ser 340 | Arg | Leu | Lys | Thr | Leu 345 | Gly | Ile | Gln | Val | Thr 350 | Lys | Val |
| Tyr | Glu | Glu 355 | Glu | Trp | Ser | Tyr | Ile 360 | Pro | Val | Gly | Gly | Ser 365 | Leu | Pro | Asn |
| Thr | Glu 370 | Gln | Lys | Asn | Leu | Ala 375 | Phe | Gly | Ala | Ala | Ala 380 | Ser | Met | Val | His |
| Pro 385 | Ala | Thr | Gly | Tyr | Ser 390 | Val | Val | Arg | Ser | Leu 395 | Ser | Glu | Ala | Pro | Lys 400 |
| Tyr | Ala | Ser | Val | Ile | Ala | Lys | Ile | Leu | Lys | Gln | Asp | Asn | Ser | Ala | Tyr |

| WO 99/63055 | |
|-------------|--|

415

405 410

Val Val Ser Gly Gln Ser Ser Ala Val Asn Ile Ser Met Gln Ala Trp 420 425

Ser Ser Leu Trp Pro Lys Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu

Phe Gly Leu Glu Leu Ile Val Gln Leu Asp Ile Glu Ala Thr Arg Thr 455

Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp Trp Gly Phe

Leu Gly Ser Ser Leu Ser Ser Phe Asp Leu Val Leu Phe Ser Met Tyr

Met Phe Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His 500 505

Leu Leu Ser Asp Pro Ser Gly Ala Val Met Val Arg Ala Tyr Leu Glu 520 525

Arg

<210> 24

<211> 1370

<212> DNA

<213> Potato

<400> 24

TAGCGGAGGA TGAGTTCAAA GATCTTGGTC TTCAAGCCTG CATTGAACAT GTTTGGCTGG 60 GATACCATTG TATATCTTGA TGATGATGAT CCTATTCTTA TTGGCCGTGC CTATGGAAGA 120 GTTAGTCGCC ATTTACTGCA CGAGGAGTTA CTCAAAAGGT GTGTGGAGGC AGGTGTTTTG 180 TATCTAAACT CGAAAGTGGA TAGGATTGTT GAGGCCACAA ATGGCCACAG TCTTGTAGAG 240 TGCGAGGGTG ATGTTGTGAT TCCCTGCAGG TTTGTGACTG TTGCATCGGG AGCAGCCTCG 300 GGGAAATTCT TGCAGTATGA GTTGGGAGGT CCTAGAGTTT CTGTTCAAAC AGCTTATGGA 360 GTGGAAGTTG AGGTCGATAA CAATCCATTT GACCCGAGCC TGATGGTTTT CATGGATTAT 420 AGAGACTATG TCAGACACGA CGCTCAATCT TTAGAAGCTA AATATCCAAC ATTTCTCTAT 480 GCCATGCCCA TGTCTCCAAC ACGAGTCTTT TTCGAGGAAA CTTGTTTGGC TTCAAAAGAT 540 GCAATGCCAT TCGATCTGTT AAAGAAAAA TTGATGTTAC GATTGAACAC CCTCGGTGTA 600 AGAATTAAAG AAATTTATGA GGAGGAATGG TCTTACATAC CAGTTGGAGG ATCTTTGCCA 660 AATACAGAAC AAAAAACACT TGCATTTGGT GCTGCTGCTA GCATGGTTCA TCCAGCCACA 720 GGTTATTCAG TCGTCAGATC ACTGTCTGAA GCTCCAAAAT GCGCCTTCGT GCTTGCAAAT 780 ATATTACGAC AAAATCATAG CAAGAATATG CTTACTAGTT CAAGTACCCC GAGTATTTCA 840 ACTCAAGCTT GGAACACTCT TTGGCCACAA GAACGAAAAC GACAAAGATC GTTTTTCCTA 900

WO 99/63055



| TTTGGACTGG | CTCTGATATT | GCAGCTGGAT | ATTGAGGGGA | TAAGGTCATT | TTTCCGCGCG | 960 |
|---|------------|------------|------------|------------|------------|------|
| TTCTTCCGTG | TGCCAAAATG | GATGTGGCAG | GGATTTCTTG | GTTCAAGTCT | TTCTTAGCAG | 1020 |
| ACCTCATGTT | ATTTGCCTTC | TACATGTTTA | TTATTGCACC | AAATGACATG | AGAAGAGGCT | 1080 |
| TAATCAGACA | TCTTTTATCT | GATCCTACTG | GTGCAACATT | GATAAGAACT | TATCTTACAT | 1140 |
| TTTAGAGTAA | ATTCCTCCTA | CAATAGTTGT | TGAAAGAGGC | CTCATTACTT | CAGATTCATA | 1200 |
| ACAGAAATCG | CGGTCTCTCG | AGGCCTTGTA | TATAACATTT | TCACTAGGTT | AATATTGCTT | 1260 |
| GAATAAGTTG | CACAGTTTCA | GTTTTTGTAT | CTGCTTCTTT | TTTGTCCAAG | ATCATGTATT | 1320 |
| GACCAATTTA | TATACATTGC | CAGTATATAT | AAATTTTATA | AAAAAAAA | | 1370 |
| <210> 25 <211> 377 <212> PRT <213> Potat | .0 | | • | | | |
| - 400 · 00 | | | | | | |

<400> 25

Asp Glu Phe Lys Asp Leu Gly Leu Gln Ala Cys Ile Glu His Val Trp

Arg Asp Thr Ile Val Tyr Leu Asp Asp Asp Pro Ile Leu Ile Gly

Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu His Glu Glu Leu Leu

Lys Arg Cys Val Glu Ala Gly Val Leu Tyr Leu Asn Ser Lys Val Asp

Arg Ile Val Glu Ala Thr Asn Gly His Ser Leu Val Glu Cys Glu Gly

Asp Val Val Ile Pro Cys Arg Phe Val Thr Val Ala Ser Gly Ala Ala

Ser Gly Lys Phe Leu Gln Tyr Glu Leu Gly Gly Pro Arg Val Ser Val 100

Gln Thr Ala Tyr Gly Val Glu Val Glu Val Asp Asn Asn Pro Phe Asp 120

Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp Tyr Val Arg His Asp 135

Ala Gln Ser Leu Glu Ala Lys Tyr Pro Thr Phe Leu Tyr Ala Met Pro 150 155

Met Ser Pro Thr Arg Val Phe Phe Glu Glu Thr Cys Leu Ala Ser Lys 170

Asp Ala Met Pro Phe Asp Leu Leu Lys Lys Leu Met Leu Arg Leu 180

Asn Thr Leu Gly Val Arg Ile Lys Glu Ile Tyr Glu Glu Glu Trp Ser 200

- WO 99/63055 Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Thr Leu 215 Ala Phe Gly Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys Cys Ala Phe Val Leu Ala Asn Ile Leu Arg Gln Asn His Ser Lys Asn Met Leu Thr Ser Ser Ser Thr Pro Ser Ile Ser Thr Gln Ala Trp Asn Thr Leu Trp Pro Gln Glu Arg Lys Arg Gln Arg Ser Phe Phe Leu Phe Gly Leu Ala Leu Ile Leu 295 Gln Leu Asp Ile Glu Gly Ile Arg Ser Phe Phe Arg Ala Phe Phe Arg 315 Val Pro Lys Met Met Trp Gly Phe Leu Gly Ser Ser Leu Ser Xaa Ala Asp Leu Met Leu Phe Ala Phe Tyr Met Phe Ile Ile Ala Pro Asn Asp 345 Met Arg Arg Gly Leu Ile Arg His Leu Leu Ser Asp Pro Thr Gly Ala 360 Thr Leu Ile Arg Thr Tyr Leu Thr Phe <210> 26
- <211> 533
- <212> PRT
- <213> Chimeric lettuce/potato
- Met Glu Cys Phe Gly Ala Arg Asn Met Thr Ala Thr Met Ala Val Phe 10
- Thr Cys Pro Arg Phe Thr Asp Cys Asn Ile Arg His Lys Phe Ser Leu
- Leu Lys Gly Arg Arg Phe Thr Asn Leu Ser Ala Ser Ser Ser Leu Arg
- Gln Ile Lys Cys Ser Ala Lys Ser Asp Arg Cys Val Val Asp Lys Gln
- Gly Ile Ser Val Ala Asp Glu Glu Asp Tyr Val Lys Ala Gly Gly Ser
- Glu Leu Phe Phe Val Gln Met Gln Arg Thr Lys Ser Met Glu Ser Gln
- Ser Lys Leu Ser Glu Lys Leu Ala Gln Ile Pro Ile Gly Asn Cys Ile
- Leu Asp Leu Val Val Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala



| | | 115 | | | | | 120 | | | | - | 125 | | | La Ĉiji |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Ala | Glu 130 | Ser | Ala | Lys | Leu | Gly 135 | Leu | Asn | Val | Glý | Leu 140 | Ile | Gly. | Pro | Asp' |
| Leu 145 | Pro | Phe | Thr | Asn | Asn 150 | Tyr | Gly | Val | Trp | Gln 155 | Asp | Glu | Phe | Ile | Gly 160 |
| Leu | Gly | Leu | Glu | Gly 165 | Cys | Ile | Glu | His | Ser 170 | Trp | Lys | Asp | Thr | Leu 175 | Val |
| Tyr | Leu | Asp | Asp 180 | Ala | Asp | Pro | Ile | Arg 185 | Ile | Gly | Arg | Ala | Tyr 190 | Gly | Arg |
| Val | His | Arg 195 | Asp | Leu | Leu | His | Glu 200 | Glu | Leu | Leu | Arg | Arg 205 | Cys | Val | Glu |
| Ser | Gly 210 | Val | Ser | Tyr | Leu | Ser 215 | Ser | Lys | Val | Glu | Arg 220 | Ile | Thr | Glu | Ala |
| Pro 225 | Asn | Gly | Tyr | Ser | Leu 230 | Ile | Glu | Cys | Glu | Gly 235 | Asn | Ile | Thr | Ile | Pro 240 |
| Cys | Arg | Leu | Ala | Thr 245 | Val | Ala | Ser | Gly | Ala 250 | Ala | Ser | Gly | Lys | Phe 255 | Leu |
| Glu | Tyr | Glu | Leu 260 | Gly | Gly | Pro | Arg | Val 265 | Ser | Val | Gln | Thr | Ala 270 | Tyr | Gly |
| Val | Glu | Val 275 | Glu | Val | Asp | Asn | Asn 280 | Pro | Phe | Asp | Pro | Ser 285 | Leu | Met | Val |
| Phe | Met 290 | Asp | Tyr | Arg | Asp | Tyr 295 | Val | Arg | His | Asp | Ala 300 | Gln | Ser | Leu | Glu |
| Ala 305 | Lys | Tyr | Pro | Thr | Phe 310 | Leu | Tyr | Ala | Met | Pro 315 | Met | Ser | Pro | Thr | Arg 320 |
| Val | Phe | Phe | Glu | Glu 325 | Thr | Cys | Leu | Ala | Ser 330 | Lys | Asp | Ala | Met | Pro 335 | Phe |
| Asp | Leu | Leu | Lys 340 | Lys | Lys | Leu | Met | Leu 345 | Arg | Leu | Asn | Thr | Leu 350 | Gly | Val |
| Arg | Ile | Lys 355 | Glu | Ile | Tyr | Glu | Glu 360 | Glu | Trp | Ser | Tyr | Ile 365 | Pro | Val | Gly |
| Gly | Ser 370 | Leu | Pro | Asn | Thr | Glu 375 | Gln | Lys | Thr | Leu | Ala 380 | Phe | Gly | Ala | Ala |
| Ala 385 | Ser | Met | Val | His | Pro 390 | Ala | Thr | Gly | Tyr | Ser 395 | Val | Val | Arg | Ser | Leu 400 |
| Ser | Glu | Ala | Pro | Lys 405 | Cys | Ala | Phe | Val | Leu 410 | Ala | Asn | Ile | Leu | Arg 415 | Gln |

Asn His Ser Lys Asn Met Leu Thr Ser Ser Ser Thr Pro Ser Ile Ser 420 425 430

Thr Gln Ala Trp Asn Thr Leu Trp Pro Gln Glu Arg Lys Arg Gln Arg

445



- Ser Phe Phe Leu Phe Gly Leu Ala Leu Ile Leu Gln Leu Asp Ile Glu
- Gly Ile Arg Ser Phe Phe Arg Ala Phe Phe Arg Val Pro Lys Trp Met 465 470 475 480

455

- Trp Gln Gly Phe Leu Gly Ser Ser Leu Ser Xaa Ala Asp Leu Met Leu 485 490 495
- Phe Ala Phe Tyr Met Phe Ile Ile Ala Pro Asn Asp Met Arg Arg Gly 500 505 510
- Leu Ile Arg His Leu Leu Ser Asp Pro Thr Gly Ala Thr Leu Ile Arg 515 520 525
- Thr Tyr Leu Thr Phe 530
- <210> 27
- <211> 374
- <212> PRT
- <213> Arabidopsis thaliana
- <400> 27
- Glu Asp Glu Phe Asn Asp Leu Gly Leu Gln Lys Cys Ile Glu His Val 1 5 10 15
- Trp Arg Glu Thr Ile Val Tyr Leu Asp Asp Lys Pro Ile Thr Ile 20 25 30
- Gly Arg Ala Tyr Gly Arg Val Ser Arg Arg Leu Leu His Glu Glu Leu 35 40 45
- Leu Arg Arg Cys Val Glu Ser Gly Val Ser Tyr Leu Ser Ser Lys Val 50 60
- Asp Ser Ile Thr Glu Ala Ser Asp Gly Leu Arg Leu Val Ala Cys Asp 65 70 75 80
- Asp Asn Asn Val Ile Pro Cys Arg Leu Ala Thr Val Ala Ser Gly Ala 85 90 95
- Ala Ser Gly Lys Leu Leu Gln Tyr Glu Val Gly Gly Pro Arg Val Cys 100 105 110
- Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Glu Asn Ser Pro Tyr 115 120 125
- Asp Pro Asp Gln Met Val Phe Met Asp Tyr Arg Asp Tyr Thr Asn Glu 130 135 140
- Lys Val Arg Ser Leu Glu Ala Glu Tyr Pro Thr Phe Leu Tyr Ala Met 145 150 155 160
- Pro Met Thr Lys Ser Arg Leu Phe Phe Glu Glu Thr Cys Leu Ala Ser
- Lys Asp Val Met Pro Phe Asp Leu Leu Lys Thr Lys Leu Met Leu Arg
- Leu Asp Thr Leu Gly Ile Arg Ile Leu Lys Thr Tyr Glu Glu Glu Trp

| wo | 99/63055 | 8 |
|------|----------|---|
| ** 0 | 77103033 | |

| | | 195 | | | | 200 | | | | 205 | | |
|---|---|--------------|---|---------|-----|-----|---|---|---|------|--|--|
| C | m | - 1 . | 5 | - 1 | - 1 | _ | _ | _ | _ | | | |

Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Asn 210 215 220

Leu Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr 225 230 235 240

Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys Tyr Ala Ser Val Ile 245 250 255

Ala Glu Ile Leu Arg Glu Glu Thr Thr Lys Gln Ile Asn Ser Asn Ile 260 265 270

Ser Arg Gln Ala Trp Asp Thr Leu Trp Pro Pro Glu Arg Lys Arg Gln 275 280 285

Arg Ala Phe Phe Leu Phe Gly Leu Ala Leu Ile Val Gln Phe Asp Thr 290 295 300

Glu Gly Ile Arg Ser Phe Phe Arg Thr Phe Phe Arg Leu Pro Lys Trp 305 310 315 320

Met Trp Gln Gly Phe Leu Gly Ser Thr Leu Thr Ser Gly Asp Leu Val
325 330 335

Leu Phe Ala Leu Tyr Met Phe Val Ile Ser Pro Asn Asn Leu Arg Lys 340 345 350

Gly Leu Ile Asn His Leu Ile Ser Asp Pro Thr Gly Ala Thr Met Ile 355 360 365

Lys Thr Tyr Leu Lys Val 370

<210> 28

<211> 1002

<212> DNA

<213> Adonis palaestina

<400> 28

ATTCATCTTC AGCAGCGCTG TCGTACTCTT TCTATATCTT CTTCCATCAC TAACAGTAGT 60 CGCCGACGGT TGAATCGGCT ATTCGCCTCA ACGTCAACTA TGGGTGAAGT CACTGATGCT 120 GGAATGGATG CTGTTCAGAA GCGGCTCATG TTCGACGACG AATGTATTTT GGTGGATGAG 180 AATGACAAGG TCGTCGGGCA TGATTCCAAA TACAACTGTC ATTTGATGGA AAAGATAGAG 240 GCAGAAAATT TGCTTCACAG AGCCTTCAGT GTTTTCTTGT TCAACTCAAA ATATGAATTG 300 CTTCTTCAGC AACGATCCGC CACAAAGGTA ACATTCCCGC TCGTATGGAC AAACACATGT 360 TGCAGTCATC CTCTCTTCG TGATTCCGAG CTCATAGAAG AAAATTATCT CGGTGTACGA 420 AACGCTGCAC AAAGAAAGCT TTTAGACGAG CTAGGCATTC CAGCTGAAGA TGTCCCAGTT 480 GATGAATTTA CTCCTCTTGG TCGCATTCTT TACAAAGCTC CATCTGACGG CAAATGGGGA 540 GAGCACGAAT TGGACTATCT CCTATTTATT GTCCGAGATG TGAAATACGA TCCAAACCCA 600 WO 99/63055 PCT/US99/12121

| GCTGATGCTG GTGAAGAGGG ACTCAAGTTG TCTCCTTGGT TTAGATTGGT TGTTGATAAC TTTTTTGTTCA AGTGGTGGGA TCATGTAGAG CAGGGTACGA TTAAGGAAGT TGCTGACATG AAAACTATCC ACAAGTTGAC TTAAGAGGAC TTCTCTCCTC TGTTCTACTA TTTGTTTTTT GCTACAATAA GTGGGTGGT ATAAGCAGTT TTTCTGTTTT CTTTAATTTA TGGCTTTTGA ATTTGCCTCC ATGTTGAACT TGTAACATAT TTAGACAAAT ATGAGACCTT GTAAGTTGAA TTTGAGGCTG AATTTATATT TTTGGGAACA TAATAATGTT AA <210> 29 <211> 1271 | | | | | | | |
|--|-------------------------|-------------------|--------------|------------|----------------------|-----------------|------|
| TTTTTGTTCA AGTGGTGGA TCATGTAGAG CAGGGTACGA TTAAGGAAGT TGCTGACATG AAAACTATCC ACAAGTTGAC TTAAGAGGAC TTCTCTCTC TGTTCTACTA TTTGTTTTTT GCTACAATAA GTGGGTGGTG ATAAGCAGTT TTTCTGTTTT CTTTAATTTA TGGCTTTTGA ATTTGCCTCG ATGTGAACT TGTAACATAT TTAGACAAAT ATGAGACCTT GTAAGTTGAA ATTTGCCTCG ATGTGAACT TGTAACATAT TTAGACAAAAT ATGAGACCTT GTAAGTTGAA **C210 | GATGAAGTTG | CTGATGCTAA | GTATGTTAAT | CGCGAGGAGT | TGAGAGAGAT | ACTGAGAAAA | 660 |
| AAAACTATCC ACAAGTTGAC TTAAGAGGAC TTCTCTCCTC TGTTCTACTA TTTGTTTTTT GCTACAATAA GTGGGTGGTG ATAAGCAGTT TTTCTGTTTT CTTTAATTTA TGGCTTTTGA ATTTGCCTCG ATGTTGAACT TGTAACATAT TTAGACAAAT ATGAGACCTT GTAAGTTGAA ATTTGAGGCTG AATTTATATT TTTGGGAACA TAATAATGTT AA 10 210> 29 | GCTGATGCTG | GTGAAGAGGG | ACTCAAGTTG | TCTCCTTGGT | TTAGATTGGT | TGTTGATAAC | 720 |
| ATTTGCCTCG ATGTGGACT TGTAACATAT TTAGACAAAT ATGAGACCTT GTAAGTTGAA ATTTGCCTCG ATGTTGAACT TGTAACATAT TTAGACAAAT ATGAGACCTT GTAAGTTGAA TTTGAGGGCTG AATTTATATT TTTGGGAACA TAATAATGTT AA (210> 29 (211> 1271 (212> DNA (213> Adonis palaestina (400> 29 TTTTAAAGCT CTTTCGCTCC ACCACCATCA AAGCCAGCCA AATTTCTCTG TACAAAAGTT TAAACCAGTA ATTCTCGGTT TGGCCCCTCC ATATCGGAAT CCTTGTTTAC GATACGCATC AAAAACACCG CTTTGGGCTT TGCCCCCCCC ATATCGGAAT CCTTGTTTAC GATACGCATC TAAACCAGTA ATTCTCGGTT TTAATTTGTT TCCTAAATTA GGCCCCTTC CGGAATCCCG AGAATTATCG CTTCGTGTC TTCTCCCGCT GTTCATCTTC AGCAGCGTTG TACGTACTCT TTCTATTTCT TCTTCCATCA CTAACAGTC TGGCCAGGG TTGAATCGGC TGTTCGCCTC AACGTCGACT ATGGGTGAAG TCGCTGATGC TGGTATGGA GCCGTCCAGA AGCGGTTAT ATACAACTGT CATTTGATGG AAAAGATCAG GGCAGAAAAC TTGCTTCACA GAGCCTTCAG ATACAACTGT CATTTGATGG AAAAGATAGA GGCAGAAAAC TTGCTTCACA GAGCCTTCAG ACATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCCTCTCC GTGATTCCGA ACCATCAGAAA GAAAATTTC TCGGGGTACG AAACACCTG TGGACCATC CCCCCTCTCC GTGATTCCGA ACCATCAGAAA GAAAATTTC TCGGGGTACG AAACCCTG TAGAGAACC TTTTAGACGA ACCATCATAGAA GAAAATTTC TCGGGGTACG AAACCCTG TGAGACCAT CCCCCTCTTC GTGATTCCGA ACCATCATAGAA GAAAATTTC TCGGGGTACG AAACCCTG CAAAGGAAGC TTTTAGACGA ACCTCATAGAA GAAAATTTC TCGGGGTACG AAACCCTG TGAGACCAT CCCCCTCTTC GTGATTCCGA ACCTCATAGAA GAAAATTTC TCGGGGTACG AAACCCTG CAAAGGAAGC TTTTAGACGA ACCTCATAGAA GAAAATTTC TCGGGGTACG AAACCCTG CCCCCTCTC GTGATTCCGA ACCTCATAGAA GAAAATTTC TCGGGGTACG AAACCCTG CAAAGGAAGC TTTTAGACGA ACCTCATAGAA GAAAATTTC TCGGGGTACG AAACCCTG CCCCCTCTTC GTCGCATTCT TTACAAAGCT CCACCTGAAGA ACCTACAACCC AGATGAAGTT GCTGGACCTA AGTACCTTAA TGCCGAGGAA TTAGAATACG ATCCAAACCC AGATGAAGTT GCTGGACCTA AGTACCTTAA CGCGAGGGAA ATTAAGGACA TACTGAGAAA AGCTGACCA GGTGAAGAGG GAATAAAGTT TCGCCGAGGAA ATTAAGGAC TCCCCACAT GAAAACCTAC CACAAGTTGA CTTAAGAGAA AGCTCCTTAA GTTCACATT TTGGTTTTTG CTTCAATAG TGGATGGAA TGAACCATTT TTATAGCTTCC TTTAATTTTG GCTTTCAAT TTGGTTTTTG TTGTTTTATG TGTTGAACTT TTAAACATATT 112 | TTTTTGTTCA | AGTGGTGGGA | TCATGTAGAG | CAGGGTACGA | TTAAGGAAGT | TGCTGACATG | 780 |
| ATTTGCCTCG ATGTTGAACT TGTAACATAT TTAGACAAAT ATGAGACCTT GTAAGTTGAA TTTGAGGCTG AATTTATATT TTTGGGAACA TAATAATGTT AA 210 > 29 <211 > 1271 <212 > DNA <213 > Adonis palaestina <400 > 29 TTTTAAAGCT CTTTGGCTCC ACCACCATCA AAGCCAGCCA AATTTCTCTG TACAAAAGTT AAAAACACCG CTTTGGGCTT TGGCCCCTCC ATATCGGAAT CCTTGTTTAC GATACGCATC TAAACCAGTA ATTCTCGGTT TTAATTTGTT TCCTAAATTA GGCCCCTTC CGGAATCCCG AGAATTATGT CGTCGATCAG GATTAATCCT TTATATAGTA TCTTCCCAC CACCACTAAA ACATTATCAG CTTCGTGTC TTCTCCCGCT GTCATCTTC AGCAGCGTTG TACGTACTCT TTCTATTTCT TCTTCCATCA CTAACAGTCC TGGCTAGGAT GCCGTCCAGA AGCGGCTTAT AACGTCGACT ATGGGTGAAG TCGCTGATGC TGGTATGGAT GCCGTCCAGA AGCGGCTTAT ATACAACTG CATTTGATG AAAAGATAGA GGCAGAAAAC TTGCTTCACA GAGCCTTCAG ATTCTAATTCTA TTCAACTCAA AATACGAGTT GCTTCTTCACA GACCACTCAG ACACTTCCCG CTCGTATGGA CAAACACCTG TTGCAGCAAAC CCCCTCTTCC GTGATTCCGA ACCTCATAGAA GAAAATTTC TCGGGGTACG AAACACCTG TTGCAGCCAC CAACAGAAGGT AACATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ACCTCATAGAA GAAAATTTC TCGGGGTACG AAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ACCTCATAGAA GAAAATTTC TCGGGGTACG AAACACCTG TGCAGCACA CTCCCTCTTCC GTGATTCCGA ACCTCATAGAA GAAAATTTC TCGGGGTACG AAACACCTG TGAGAACAC TTGCTTCACA GACCTTCTC TTACAAAAGCT CCACCTGAAG ACGTACCAGT TGATGAATTC ACCCCTCTTC GTCGCATTCT TTACAAAGCT CCACCTGAAG ACGTACCAGT TGATGAATTC ACCCCTCTTC GTCGCATTCT TTACAAAGCT CCACCTGAAG ACGTACCAGT TGATGAATTC ACCCCTCTTC GTCGCATTCT TTACAAAGCT CCACCTGAAG ACGTACCAGT TGATGAACTT GCTGACCTTA AGTACGTTAA TCGCCGAGGAG TTGAAAAGAGA TACTGAGAAA AGCTGATGCA GGTGAACGAG CTGGACTTAT TCGCCGAGGAG TTGAAAAGAG TACTGAGAAA AGCTGATGCA GGTGAACACGT AGTACGTTAA TCGCCGAGGAG TTGAAAAGAGA TACTGAGAAA AGCTGATGCA GGTGAACGAG GAATAAAGGTT TCGCCGAGGAG TTGAAAAGAGA TACTGAGAAA AGCTGATGC AAAGCTGAG GAAAAACATT TCGCCGAGGAA ATTAAGGAC TCGCCCCACAT GAAACCTC CACAAGTTG CTTAAGAGAA TCGCCGAGGAG TTGAAAAGAGA TACTGAGAAA AGCTGATGC GAAAACTTC CACAAGTTG CTTAAGAGAA TTATCCTTC TTTAATTTG GCTTTCAAT TTGGTTTTG CTTCAATAAG TGGTGAGCTT TAAGAGAATT TTATCCTTC TTTAATTTTG GCTTTCAAT TTGGTTTTAG TTTTCAATTTTG TTTAAAAACTTT TTGGTTTTAG | AAAACTATCC | ACAAGTTGAC | TTAAGAGGAC | TTCTCTCCTC | TGTTCTACTA | TTTGTTTTT | 840 |
| **210> 29 **211> 1271 **212> DNA **213> Adonis palaestina ***400> 29 **TTTTAAAGCT CTTTCGCTCC ACCACCATCA AAGCCAGCCA AATTTCTCTG TACAAAAGTT ***TAAAACACCG CTTTGGGCTT TGGCCCCTCC ATATCGGAAT CCTTGTTTAC GATACGCATC ***TAAAACACCG CTTTGGGTT TTAATTTGTT TCCTAAATTA GGCCCCTTC CGGAATCCCG ***AGAATTATGT CGTCGATCA GATTAATCCT TTATATAGTA TCTTCTCAC CACCACTAAA ***ACACTTATCAG CTTCGTGTC TTCTCCCGCT GTTCATCTTC AGCAGCGTT TACGTACTCT ***TCTATTTCT TCTTCCATCA CTAACAGTC TCGCCGAGGG TTGAATCGGC TGTTCGCCTC ***AACGTCGACT ATGGGTGAAG TCGCTGATGC TGGTATGGAT GCCGTCCAGA AGCGCTTAT ***GTTCGACGAT GAATGTATTT TGGTGGATGA GAATGACAAG GTCGTCGGAC ATGATTCCAA ***ATACAACTGT CATTTGATGG AAAAGATAGA GGCAGAAAAC TTGCTTCACA GAGCCTTCAG ***AACATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCAT CCCCTCTTCC GTGATTCCGA ***AACATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCAT CCCCTCTTCC GTGATTCCGA ***ACCATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCAT CCCCTCTTCC GTGATTCCGA ***ACCATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ***ACCATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ***ACCATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ***ACCATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ***ACCATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ***ACCATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ***ACCATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCTG ***TTTACAAAGCT CCACCTGAAG ACGTACCAGT TGATGAATTC ACTCCTCTTG GTCGCATTCT ***TTACAAAGCT CCACCTGAAG ACGTACCAGT TGATGAATTC ACTCCTCTTG GTCGCATTCT ***TTCCCGCGAGGA TTGAAAGAGA TACTGAGAAA AGCTGATCC AAAGGAAGG GAATAAAGTT ***TTCCCGCGAGAG TTGAAAGAGA TACTGAGAAA AGCTGATCC CACAAGTTGA CTTAAGAGAA ***GTCTCCTTGG TTTAAGTTGG TTGTGGATAA CTTTTTTGTC AAGTGGTGG ATCATGTAGA ***GGGGGGAAGA ATTAAGGACG TCGCCGACAT GAAAACTATC CACAAGTTGA CTTAAGAGAA ***GTCTCTTAA GTTCAATT TTGGTTTTTG CTTCAATAAG TGGATGGTG TGAGCAGTTT ***TTATCTCCTC TTTAATTTTG GCTTTTCAAT TTGGTTTTAG TGTTGAACTT GTAACATATT TTGTTTTAATTTTG GTTTTTAGT TTTTTATTTTTTTT | GCTACAATAA | GTGGGTGGTG | ATAAGCAGTT | TTTCTGTTTT | СТТТААТТТА | TGGCTTTTGA | 900 |
| <pre><210> 29 <211> 1271 <212> DNA <213> Adonis palaestina </pre> <pre><400> 29 TTTTAAAGCT CTTTCGCTCC ACCACCATCA AAGCCAGCCA AATTTCTCTG TACAAAAGTT AAAAACACCG CTTTGGGCTT TGGCCCCTCC ATATCGGAAT CCTTGTTTAC GATACGCATC TAAACCAGTA ATTCTCGGTT TTAATTTGTT TCCTAAATTA GGCCCCTTTC CGGAATCCCG AGAATTATGT CGTCGATCAG GATTAATCCT TTATATAGTA TCTTCTCCAC CACCACTAAA ACATTATCAG CTTCGTGTC TTCTCCCGCT GTTCATCTTC AGCAGCGTTG TACGTACTCT TTCTATTTCT TCTTCCATCA CTAACAGTCC TCGCCGAGGG TTGAATCGGC TGTTCGCCTC AACGTCGACT ATGGTGAAG TCGCTGATGC TGGTATGGAT GCCGTCCAGA AGCGGCTTAT GTTCGACCGAT GAATGTATTT TGGTGGATGA GAATGACAAG GTCGTCGAGA AGCGCTTCAG ATACAACTGT CATTTGATGG AAAAGATAGA GGCAGAAAAC TTGCTTCACA GAGCCTTCAG TGTTTTCTTA TTCAACTCAA AATACGAGTT GCTTCTTCAG CAACGAAGGT GACATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ACCTCATAGAA GAAAATTTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA ACCTCATAGAA GAAAATTTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA ACCTCATAGAA GAAAATTTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA ACCTCATAGAA GAAAATTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA ACCTCATAGAA GAAAATTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA ACCTCATAGAA GAAAATTTC TCGGGGTACG AGAGCCCGA CTGGACTATC TTCTGTTTAT TGCCCGAGGAG GTGAAATACG ATCCAAACCC AGATGAAGT GCTGACGCTA AGTACGTTAA TCGCCGAGGAG TTGAAAGAGA TACTGAGAAA AGCTGATCA GCTGACGCTA AGTACGTTAA TCGCCGAGGAG TTGAAAGAGA TACTGAGAAA ACCTTGTTCTC CACAAGTTGA CTTAAGAGAA AGCTCCTTAG GTGAAATACG TCGCCGACAT GAAAACTATC CACAAGTTGA CTTAAGAGAA AGCTCCTTAG GTGAAATACG TCGCCGACAT GAAAACTATC CACAAGTTGA CTTAAGAGAA AGCTCCTTAA GTTCAACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTA TCAAACCATT TTATGCTTCC TTTAATTTTG GCTTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT 12</pre> | ATTTGCCTCG | ATGTTGAACT | TGTAACATAT | TTAGACAAAT | ATGAGACCTT | GTAAGTTGAA | 960 |
| <pre><211> 1271 <212> DNA <213> Adonis palaestina </pre> <pre><400> 29 TTTTAAAGCT CTTTCGCTCC ACCACCATCA AAGCCAGCCA AATTTCTCTG TACAAAAGTT AAAAACACCG CTTTGGGCTT TGGCCCCTCC ATATCGGAAT CCTTGTTTAC GATACGCATC TAAACCAGTA ATTCTCGGTT TTAATTTGTT TCCTAAATTA GGCCCCTTTC CGGAATCCCG AGAATTATGT CGTCGATCAG GATTAATCCT TTATATAGTA TCTTCTCCAC CACCACTAAA ACATTATCAG CTTCGTGTTC TCTCCCGCT GTTCATCTTC AGCAGCGTTG TACGTACTCT TTCTATTTCT TCTTCCATCA CTAACAGTCC TCGCCGAGGG TTGAATCGGC TGTTCGCCTC AACGTCGACT ATGGGTGAAG TCGCTGATGC TGGTATGGAT GCCGTCCAGA AGCGGCTTAT GTTCGACCGAT GAATGTATTT TGGTGGATGA GAATGACAAG GTCGTCGGAC ATGATTCCAA ATACAACTGT CATTTGATGG AAAAGATAGA GGCAGAAAAC TTGCTTCACA GAGCCTTCAG TGTTTTCTTA TTCAACTCAA AATACGAGTT GCTTCTCAG CAACGATCTG CAACGAAGGT AACATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ACCTCATAGAA GAAAATTTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA ACTCATAGAA GAAAATTTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA ACTCATAGAA GAAAATTTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA ACTCATAGAA GAAAATTTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA ACTCCATAGAA GAAAATTTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTTAGACGA ACTCCATAGAA GAAAATTTC TCGGGGTACG AGAGCACGAA CTGGACTATC TTCTGTTTAT TGCCCGAGGAG TTGAAAACACC AGATGAAGTT GCTGACGCTA AGTACGTTAA TCGCCGAGGAG TTGAAAAACC ATCCAAACCC AGATGAAGTT GCTGACGCTA AGTACGTTAA TCGCCGAGGAG TTGAAAGAGA TACTGAGAAA ACCTTGTTCTC AAGTACGTGA AGTACGTTAA TCGCCGAGGAG TTGAAAGAGA TACTGAGAAA ACCTTGTTCTC AAGTACGAG GAATAAAGTT GGCGGGGGAAG ATTAAAGGAC TCGCCGACAT GAAACCTTC CACAAGTTGA CTTAAGAGAA AGTCCCTTAA GTTCAACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAACCATTT TTATGCTTCC TTTAAATTTTG CCTTTCAAT TTGCTTTATG TGTTGAACCTT GTAACATATT 12</pre> | TTTGAGGCTG | AATTTATATT | TTTGGGAACA | TAATAATGTT | AA | | 1002 |
| AAAAACACCG CTTTGGCTCC ACCACCATCA AAGCCAGCCA AATTTCTCTG TACAAAAGTT AAAAACACCG CTTTGGGCTT TGGCCCCTCC ATATCGGAAT CCTTGTTTAC GATACGCATC TAAAACCAGTA ATTCTCGGTT TTAATTTGTT TCCTAAATTA GGCCCCTTTC CGGAATCCCG AGAATTATGT CGTCGATCAG GATTAATCCT TTATATAGTA TCTTCTCCAC CACCACTAAA ACCATTATCAG CTTCGTGTTC TTCTCCCGCT GTTCATCTTC AGCAGCGTTG TACGTACTCT TTCTATTTCT TCTTCCATCA CTAACAGTCC TCGCCGAGGG TTGAATCGGC TGTTCGCCTC AACGTCGACT ATGGGTGAAG TCGCTGATGC TGGTAGGAT GCCGTCCAGA AGCGGCTTAT GTTCGACGAT GAATGTATTT TGGTGGATGA GAATGACAAG GTCGTCGGAC ATGATTCCAA ATACAACTGT CATTTGATGG AAAAGATAGA GGCAGAAAAC TTGCTTCACA GAGCCTTCAG TGTTTTCTTA TTCAACTCAA AATACGAGTT GCTTCTTCAG CAACGAATCT CAACGAAGGT AACATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ACTCATAGAA GAAAATTTC TCGGGGTACG AAACCCTGCA CAAAGGAAGC TTTTAGACGA GCTAGGCATT CCAGCTGAAG ACGTACCAGT TGATGAATTC ACTCCTCTTG GTCGCATTCT TTACAAAGCT CCATCTGAC GAAAATGGGG AGAGCACGAA CTGGACTATC TTCTGTTTAT TGTCCGAGGAT GTGAAAATACG ATCCAAACCC AGATGAAGTT GCTGACCTAT TTCTGTTTAT TGTCCGAGGA TTGAAAGAGA TACTGAGAAA AGCTGATGCA GGTGAAGAGG GAATAAAGTT TCGCCGAGGA TTGAAAGAGA TACTGAGAAAA AGCTGATGCA GGTGAAGAGG GAATAAAGTT TCGCCGAGGA TTGAAAGAGA TACTGAGAAAA AGCTGATGCA GGTGAAGAGG GAATAAAGTT TCGCCGAGGA TTTAGAATAG TTGTGGATAA ACTTCTTTGTTC AAGTGGGG ATCATGTAGA GGAGGGGAAG ATTAAGGAC TCCCCGACAT GAAAACTATC CACAAGTTGA CTTAAGAGAA AGTCTCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTG TGAACACTATT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGGTTGAACTT GTATAACTTT TAGACCTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAACACTATT TTATAGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGGTTGAACTT GTAACACTATT TTATAGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGGTTGAACTT GTAACATATT TTATAGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGGTTGAACTT GTAACATATT TTATAGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGGTTGAACTT GTAACATATT TTATAGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT | <211> 1271 <212> DNA | is palaestir | na | | | | |
| AAAAACACCG CTTTGGGCTT TGGCCCCTCC ATATCGGAAT CCTTGTTTAC GATACGCATC TAAACCAGTA ATTCTCGGTT TTAATTTGTT TCCTAAATTA GGCCCCTTTC CGGAATCCCG AGAATTATGT CGTCGATCAG GATTAATCCT TTATATAGTA TCTTCTCCAC CACCACTAAA ACATTATCAG CTTCGTGTTC TTCTCCCGCT GTTCATCTTC AGCAGCGTTG TACGTACTCT TTCTATTTCT TCTTCCATCA CTAACAGTCC TCGCCGAGGG TTGAATCGGC TGTTCGCCTC AACGTCGACT ATGGGTGAAG TCGCTGATGC TGGTATGGAT GCCGTCCAGA AGCGGCTTAT GTTCGACGAT GAATGTATTT TGGTGGATGA GAATGACAAG GTCGTCCGAC ATGATTCCAA ATACAACTGT CATTTGATGG AAAAGATAGA GGCAGAAAAC TTGCTTCACA GAGCCTTCAG TGTTTTCTTA TTCAACTCAA AATACGAGTT GCTTCTTCAG CAACGAAGGT CAACGAAGGT AACATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ACTCATAGAA GAAAATTTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA GCTAGGCATT CCAGCTGAAG ACGTACCAGT TGATGAATTC ACTCCTCTTG GTCGCATTCT TTACAAAGCT CCATCTGAC GAAAATGGGG AGGCACGAA CTGGACTATC TTCTGTTTAT TGTCCGAGAT GTGAAATACG ATCCAAACCC AGATGAAGTT GCTGACCATA AGTACGTTAA TCGCGAGGGG TTGAAAAGAG TACTGAGAAA AGCTGATGCA GGTGAAGAGG GAATAAAGTT GCCGCAGGAG TTGAAAAGAG TACTGAGAAA AGCTGATGCA GGTGAACAGG GAATAAAGTT GCCGCAGGAG TTTAGATTGG TTGTGGATAA AGCTGATGCA GAAGGAGG GAATAAAGTT GCCGCAGGAG TTTAGATTGG TTGTGGATAA CTTTTTGTTC AAGTGGGAGA ACCTCCTTTGG GGAGGGGAAG ATTAAGGACG TCGCCCACAT GAAACCTAC CACAAGTTGA CTTAAGAGAA AGTCCCTTAA GTTCACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAGCAGTTT TTATAGCTTCC TTTAATTTTG GCTTTTTAGATTG TTGCTTTATT GTGTATAATT CACAAGTTGA CTTAAGAGAA AGTCCTCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAGCAGTTT TTATAGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGGTAGACTT GTAACATATT | | CTTTCGCTCC | מרימרנמיינים | AACCCACCCA | እ አ ምምም ርምርምር | MACAAAAC | 60 |
| TARACCAGTA ATTCTCGGTT TTAATTTGTT TCCTAAATTA GGCCCCTTTC CGGAATCCCG AGAATTATGT CGTCGATCAG GATTAATCCT TTATATAGTA TCTTCTCCAC CACCACTAAA ACATTATCAG CTTCGTGTTC TTCTCCCGCT GTTCATCTTC AGCAGCGTTG TACGTACTCT TTCTATTTCT TCTTCCATCA CTAACAGTCC TCGCCGAGGG TTGAATCGGC TGTTCGCCTC AACGTCGACT ATGGGTGAAG TCGCTGATGC TGGTATGGAT GCCGTCCAGA AGCGGCTTAT GTTCGACGAT GAATGTATTT TGGTGGATGA GAATGACAAG GTCGTCGGAC ATGATTCCAA ATACAACTGT CATTTGATGG AAAAGATAGA GGCAGAAAAC TTGCTTCACA GAGCCTTCAG ACCATTCCAA AATACGAGTT GCTTCTCAG CAACGATCTG CAACGAAGGT AACATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ACTCATAGAA GAAAATTTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA GCTAGGCATT CCAGCTGAAG ACGTACCAGT TGATGAATTC ACTCCTCTTG GTCGCATTCT TTACAAAGCT CCATCTGACG GAAAATGGGG AGAGCACGAA CTGGACTATC TTCTGTTTAT TGTCCCGAGGAT GTGAAATACG ATCCAAACCC AGATGAAGTT GCTGACGCTA AGTACCGTTAA TCGCCGAGGAG TTGAAAGAGA TACTGAGAAA AGCTGATGCA GGTGAAGAGG GAATAAAGTT TCGCCGAGGAG TTGAAAGAGA TACTGAGAAAA AGCTGATCC AGAGGAGGG GAATAAAGTT TCGCGAGGGGAAG ATTAAAGGAC TCGCCGACAT GAAAACTAC CACAAGTTGA CTTAAGAGGAA AGCTCCTTGG TTTAGATTGG TTGTGGATAA CTTTTTGTTC AAGTGGTGG ATCATGTAGA AGTCTCCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTG TGAACATATT 11 TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGTTGAACCTT GTAACATATT 11 TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTTAT TGGTTTTATG TGTTGAACATTT 11 TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTTATG TGTTGAACTT GTAACATATT 11 TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTTATG TGTTGAACTT GTAACATATT 11 TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTTATG TGTTGAACTT GTAACATATT 11 | | | | | | | 60 |
| AGAATTATGT CGTCGATCAG GATTAATCCT TTATATAGTA TCTTCTCCAC CACCACTAAA ACATTATCAG CTTCGTGTTC TTCTCCCGCT GTTCATCTTC AGCAGCGTTG TACGTACTCT TTCTATTTCT TCTTCCATCA CTAACAGTCC TCGCCGAGGG TTGAATCGGC TGTTCGCCTC AACGTCGACT ATGGGTGAAG TCGCTGATGC TGGTATGGAT GCCGTCCAGA AGCGGCTTAT GTTCGACGAT GAATGTATTT TGGTGGATGA GAATGACAAG GTCGTCGGAC ATGATTCCAA ATACAACTGT CATTTGATGG AAAAGATAGA GGCAGAAAAC TTGCTTCACA GAGCCTTCAG TGTTTTCTTA TTCAACTCAA AATACGAGTT GCTTCTCAG CAACGATCTG CAACGAAGGT AACATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ACTCATAGAA GAAAATTTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA GCTAGGCATT CCAGCTGAAG ACGTACCAGT TGATGAATTC ACTCCTCTTG GTCGCATTCT TTACAAAGCT CCATCTGACG GAAAATGGGG AGAGCACGAA CTGGACTATC TTCTGTTTAT TGTCCGAGGAT GTGAAATACG ATCCAAACCC AGATGAAGTT GCTGAACTAC TTCTGTTTAT TCGCCGAGGAG TTGAAAGAGA TACTGAGAAA AGCTGATGCA GGTGAAGAGG GAATAAAGTT GTCCCGAGGAG TTGAAAGAGA TACTGAGAAA CTTTTTTGTTC AAGTGGTGG ATCATGTAGA GGAGGGGAAG ATTAAGGACG TCGCCGACAT GAAAACTATC CACAAGTTGA CTTAAGAGAA AGTCTCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAGCAGTTT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT | | | | | | | 120 |
| ACATTATCAG CTTCGTGTTC TTCTCCCGCT GTTCATCTTC AGCAGCGTTG TACGTACTCT TTCTATTTCT TCTTCCATCA CTAACAGTCC TCGCCGAGGG TTGAATCGGC TGTTCGCCTC AACGTCGACT ATGGGTGAAG TCGCTGATGC TGGTATGGAT GCCGTCCAGA AGCGGCTTAT GTTCGACGAT GAATGTATTT TGGTGGATGA GAATGACAAG GTCGTCGGAC ATGATTCCAA ATACAACTGT CATTTGATGG AAAAGATAGA GGCAGAAAAC TTGCTTCACA GAGCCTTCAG TGTTTTCTTA TTCAACTCAA AATACGAGTT GCTTCTCAG CAACGAAGGT AACATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ACTCATAGAA GAAAATTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA GCTAGGCATT CCAGCTGAAG ACGTACCAGT TGATGAATTC ACTCCTCTTG GTCGCATTCT TTACAAAGCT CCATCTGACG GAAAATGGGG AGAGCACGAA CTGGACTATC TTCTGTTTAT TGTCCGAGGAT GTGAAATACG ATCCAAACCC AGATGAAGTT GCTGACGCTA AGTACGTTAA TCGCCGAGGAG TTGAAAGAGA TACTGAGAAA AGCTGATGCA GAGGGGGGG ATCATGTAGA GGAGGGGAAG ATTAAGGACG TCGCCGACAT GAAAACTATC CACAAGTGA CTTAAGAGAA AGTCTCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTG TGAACCATTT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGTTGAACCTT GTAAACATTT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTTATG TGTTGAACCTT GTAAACATTT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTTATG TGTTGAACCTT GTAAACATTT | | | | | | | 180 |
| ACCOTCATA TOTACATCA CTAACAGTCC TCGCCGAGGG TTGAATCGGC TGTTCGCCTC AACGTCGACT ATGGGTGAAG TCGCTGATGC TGGTATGGAT GCCGTCCAGA AGCGGCTTAT GTTCGACGAT GAATGTATTT TGGTGGATGA GAATGACAAG GTCGTCGGAC ATGATTCCAA ATACAACTGT CATTTGATGG AAAAGATAGA GGCAGAAAAC TTGCTTCACA GAGCCTTCAG TGTTTTCTTA TTCAACTCAA AATACGAGTT GCTTCTTCAG CAACGATCTG CAACGAAGGT AACATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ACTCATAGAA GAAAATTTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA GCTAGGCATT CCAGCTGAAG ACGTACCAGT TGATGAATTC ACTCCTCTTG GTCGCATTCT TTACAAAGCT CCATCTGACG GAAAATGGGG AGAGCACGAA CTGGACTATC TTCTGTTTAT TGTCCGAGGAT GTGAAATACG ATCCAAACCC AGATGAAGTT GCTGACGCTA AGTACGTTAA TCGCCGAGGAG TTGAAAGAGA TACTGAGAAA AGCTGATGCA GGTGAAGAGG GAATAAAGTT GTCTCCTTGG TTTAGATTGG TTGTGGATAA CTTTTTGTTC AAGTGGTGGG ATCATGTAGA AGTCTCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAGCAGTTT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTTATG TGTTGAACTT GTAACATATT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTTATG TGTTGAACTT GTAACATATT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTTATG TGTTGAACTT GTAACATATT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTTATG TGTTGAACTT GTAACATATT TTATGCTTCC TTTTAATTTTG GCTTTTCAAT TTGCTTTTATG TGTTGAACTT GTAACATATT TTATGCTTCC TTTTAATTTTG GCTTTTCAAT TTGCTTTTATG TGTTGAACTT GTAACATATT | | | | | | | 240 |
| AACGTCGACT ATGGGTGAAG TCGCTGATGC TGGTATGGAT GCCGTCCAGA AGCGGCTTAT GTTCGACGAT GAATGTATTT TGGTGGATGA GAATGACAAG GTCGTCGGAC ATGATTCCAA ATACAACTGT CATTTGATGG AAAAGATAGA GGCAGAAAAC TTGCTTCACA GAGCCTTCAG TGTTTTCTTA TTCAACTCAA AATACGAGTT GCTTCTCAG CAACGATCTG CAACGAAGGT AACATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ACTCATAGAA GAAAATTTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA GCTAGGCATT CCAGCTGAAG ACGTACCAGT TGATGAATTC ACTCCTCTTG GTCGCATTCT TTACAAAGCT CCATCTGACG GAAAATGGGG AGAGCACGAA CTGGACTATC TTCTGTTTAT TGTCCGAGGAT GTGAAATACG ATCCAAACCC AGATGAAGTT GCTGACGCTA AGTACGTTAA TCGCGAGGAG TTGAAAGAGA TACTGAGAAA AGCTGATGCA GGTGAAGAGG GAATAAAGTT GTCTCCTTGG TTTAGATTGG TTGTGGATAA CTTTTTGTTC AAGTGGTGG ATCATGTAGA GGAGGGGAAG ATTAAGGACG TCGCCGACAT GAAAACTATC CACAAGTTGA CTTAAGAGAA AGTCTCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAGCAGTTT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGTTGAACCTT GTAACATATT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTTATG TGTTGAACCTT GTAACATATT | | | | | | | 300 |
| GTTCGACGAT GAATGTATTT TGGTGGATGA GAATGACAAG GTCGTCGGAC ATGATTCCAA ATACAACTGT CATTTGATGG AAAAGATAGA GGCAGAAAAC TTGCTTCACA GAGCCTTCAG TGTTTTCTTA TTCAACTCAA AATACGAGTT GCTTCTTCAG CAACGATCTG CAACGAAGGT AACATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ACTCATAGAA GAAAATTTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA GCTAGGCATT CCAGCTGAAG ACGTACCAGT TGATGAATTC ACTCCTCTTG GTCGCATTCT TTACAAAGCT CCATCTGACG GAAAATGGGG AGAGCACGAA CTGGACTATC TTCTGTTTAT TGTCCGAGGAT GTGAAATACG ATCCAAACCC AGATGAAGTT GCTGACGCTA AGTACGTTAA TCGCGAGGAG TTGAAAAGAGA TACTGAGAAA AGCTGATGCA GGTGAAGAGG GAATAAAGTT GTCTCCTTGG TTTAGATTGG TTGTGGATAA CTTTTTGTTC AAGTGGTGGA ATCATGTAGA GGAGGGGAAG ATTAAGGACG TCGCCGACAT GAAAACTATC CACAAGTTGA CTTAAGAGAA AGTCTCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAGCAGTTT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT | | | | | | | 360 |
| ATACAACTGT CATTTGATGG AAAAGATAGA GGCAGAAAAC TTGCTTCACA GAGCCTTCAG TGTTTTCTTA TTCAACTCAA AATACGAGTT GCTTCTTCAG CAACGATCTG CAACGAAGGT AACATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ACTCATAGAA GAAAATTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA GCTAGGCATT CCAGCTGAAG ACGTACCAGT TGATGAATTC ACTCCTCTTG GTCGCATTCT TTACAAAGCT CCATCTGACG GAAAATGGGG AGAGCACGAA CTGGACTATC TTCTGTTTAT TGTCCGAGAT GTGAAATACG ATCCAAACCC AGATGAAGTT GCTGACGCTA AGTACGTTAA TCGCGAGGAG TTGAAAGAGA TACTGAGAAA AGCTGATGCA GGTGAAGAGG GAATAAAGTT GTCTCCTTGG TTTAGATTGG TTGTGGATAA CTTTTTGTTC AAGTGGTGG ATCATGTAGA AGTCTCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAGCAGTTT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATTG TGTTGAACTT GTAACATATT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTTATG TGTTGAACTT GTAACATATT | | | | | | | 420 |
| TGTTTCTTA TTCAACTCAA AATACGAGTT GCTTCTCAG CAACGATCTG CAACGAAGGT AACATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ACTCATAGAA GAAAATTTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA GCTAGGCATT CCAGCTGAAG ACGTACCAGT TGATGAATTC ACTCCTCTTG GTCGCATTCT TTACAAAGCT CCATCTGACG GAAAATGGGG AGAGCACGAA CTGGACTATC TTCTGTTTAT TGTCCGAGAT GTGAAATACG ATCCAAACCC AGATGAAGTT GCTGACGCTA AGTACGTTAA TCGCGAGGAG TTGAAAGAGA TACTGAGAAA AGCTGATGCA GGTGAAGAGG GAATAAAGTT GTCTCCTTGG TTTAGATTGG TTGTGGATAA CTTTTTGTTC AAGTGGTGG ATCATGTAGA GGAGGGGAAG ATTAAGGACG TCGCCGACAT GAAAACTATC CACAAGTTGA CTTAAGAGAA AGTCTCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAGCAGTTT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGTTGAACCTT GTAACATATT 12 | | | | | | | 480 |
| AACATTCCCG CTCGTATGGA CAAACACCTG TTGCAGCCAT CCCCTCTTCC GTGATTCCGA ACTCATAGAA GAAAATTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA GCTAGGCATT CCAGCTGAAG ACGTACCAGT TGATGAATTC ACTCCTCTTG GTCGCATTCT TTACAAAGCT CCATCTGACG GAAAATGGGG AGAGCACGAA CTGGACTATC TTCTGTTTAT TGTCCGAGAT GTGAAATACG ATCCAAACCC AGATGAAGTT GCTGACGCTA AGTACGTTAA TCGCGAGGAG TTGAAAGAGA TACTGAGAAA AGCTGATGCA GGTGAAGAGG GAATAAAGTT GTCTCCTTGG TTTAGATTGG TTGTGGATAA CTTTTTGTTC AAGTGGTGG ATCATGTAGA GGAGGGGAAG ATTAAGGACG TCGCCGACAT GAAAACTATC CACAAGTTGA CTTAAGAGAA AGTCTCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAGCAGTTT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT | | | | | | , | 540 |
| ACTCATAGAA GAAAATTTC TCGGGGTACG AAACGCTGCA CAAAGGAAGC TTTTAGACGA GCTAGGCATT CCAGCTGAAG ACGTACCAGT TGATGAATTC ACTCCTCTTG GTCGCATTCT TTACAAAGCT CCATCTGACG GAAAATGGGG AGAGCACGAA CTGGACTATC TTCTGTTTAT TGTCCGAGAT GTGAAATACG ATCCAAACCC AGATGAAGTT GCTGACGCTA AGTACGTTAA TCGCGAGGAG TTGAAAAGAGA TACTGAGAAA AGCTGATGCA GGTGAAGAGG GAATAAAGTT GTCTCCTTGG TTTAGATTGG TTGTGGATAA CTTTTTGTTC AAGTGGTGGG ATCATGTAGA GGAGGGGAAG ATTAAGGACG TCGCCGACAT GAAAACTATC CACAAGTTGA CTTAAGAGAA AGTCTCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAGCAGTTT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT | | | | | | | 600 |
| GCTAGGCATT CCAGCTGAAG ACGTACCAGT TGATGAATTC ACTCCTCTTG GTCGCATTCT TTACAAAGCT CCATCTGACG GAAAATGGGG AGAGCACGAA CTGGACTATC TTCTGTTTAT TGTCCGAGAT GTGAAATACG ATCCAAACCC AGATGAAGTT GCTGACGCTA AGTACGTTAA TCGCGAGGAG TTGAAAGAGA TACTGAGAAA AGCTGATGCA GGTGAAGAGG GAATAAAGTT GTCTCCTTGG TTTAGATTGG TTGTGGATAA CTTTTTGTTC AAGTGGTGGG ATCATGTAGA GGAGGGGAAG ATTAAGGACG TCGCCGACAT GAAAACTATC CACAAGTTGA CTTAAGAGAA AGTCTCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAGCAGTTT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT | | | | | | | 660 |
| TTACAAAGCT CCATCTGACG GAAAATGGGG AGAGCACGAA CTGGACTATC TTCTGTTTAT TGTCCGAGAT GTGAAATACG ATCCAAACCC AGATGAAGTT GCTGACGCTA AGTACGTTAA TCGCGAGGAG TTGAAAGAGA TACTGAGAAA AGCTGATGCA GGTGAAGAGG GAATAAAGTT GTCTCCTTGG TTTAGATTGG TTGTGGATAA CTTTTTGTTC AAGTGGTGGG ATCATGTAGA GGAGGGGAAG ATTAAGGACG TCGCCGACAT GAAAACTATC CACAAGTTGA CTTAAGAGAA AGTCTCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAGCAGTTT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT | | | | | | | 720 |
| TGTCCGAGAT GTGAAATACG ATCCAAACCC AGATGAAGTT GCTGACGCTA AGTACGTTAA TCGCGAGGAG TTGAAAGAGA TACTGAGAAA AGCTGATGCA GGTGAAGAGG GAATAAAGTT GTCTCCTTGG TTTAGATTGG TTGTGGATAA CTTTTTGTTC AAGTGGTGGG ATCATGTAGA GGAGGGGAAG ATTAAGGACG TCGCCGACAT GAAAACTATC CACAAGTTGA CTTAAGAGAA AGTCTCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAGCAGTTT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT | | | | * | | | 780 |
| TCGCGAGGAG TTGAAAGAGA TACTGAGAAA AGCTGATGCA GGTGAAGAG GAATAAAGTT GTCTCCTTGG TTTAGATTGG TTGTGGATAA CTTTTTGTTC AAGTGGTGGG ATCATGTAGA GGAGGGGAAG ATTAAGGACG TCGCCGACAT GAAAACTATC CACAAGTTGA CTTAAGAGAA AGTCTCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAGCAGTTT TTATGCTTCC TTTAATTTTG GCTTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT | | | | | , | | 840 |
| GTCTCCTTGG TTTAGATTGG TTGTGGATAA CTTTTTGTTC AAGTGGTGGG ATCATGTAGA 100 GGAGGGGAAG ATTAAGGACG TCGCCGACAT GAAAACTATC CACAAGTTGA CTTAAGAGAA 100 AGTCTCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAGCAGTTT 110 ATTATGCTTCC TTTAATTTTG GCTTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT 120 AGTCTCTCC TTTAATTTTG GCTTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT 120 AGTCTCTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT 120 AGTCTCTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT 120 AGTCTCTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT 120 AGTCTCTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT 120 AGTCTCTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT 120 AGTCTCTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT 120 AGTCTCTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT 120 AGTCTCTCAAT AGTCTCTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT 120 AGTCTCTCAATATT 120 AGTCTCTCAATATT 120 AGTCTCTCAATATATATG TGTTGAACTT GTAACATATT 120 AGTCTCTCAATATATG TGTTGAACTT GTAACATATT 120 AGTCTCTCAATATG TGTTGAACTT GTAACATATT 120 AGTCTCTCAATATG TGTTGAACTT GTAACATATT 120 AGTCTCTCAATATG TGTTGAACTT GTAACATATT 120 AGTCTCTCAATATG TGTTGAACATATT 120 AGTCTCAATATG TGTTGAACATATT 120 AGTCTCAATATG TGTTGAACATATT 120 AGTCTCAATATG TGTTGAACATATT 120 AGTCTCAATATG TGTTGAACATATT 120 AGTCTCAATATG AGTCAATATG AGTCAATA | | | | | | | 900 |
| GGAGGGAAG ATTAAGGACG TCGCCGACAT GAAAACTATC CACAAGTTGA CTTAAGAGAA AGTCTCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAGCAGTTT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT 12 | | | | | | | 960 |
| AGTCTCTTAA GTTCTACTAT TTGGTTTTTG CTTCAATAAG TGGATGGTGA TGAGCAGTTT TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT 12 | | | | | | | 1020 |
| TTATGCTTCC TTTAATTTTG GCTTTTCAAT TTGCTTTATG TGTTGAACTT GTAACATATT 12 | | | | | | | 1080 |
| | | | | | | | 1140 |
| TAGTCAAATA TGAGACCTTG TGAGTTGAAT TTGAGGTTAT ATTTATAGTT TTGGGAACAT 12 | | | | | | | 1200 |
| | TAGTCAAATA | TGAGACCTTG | TGAGTTGAAT | TTGAGGTTAT | ATTTATAGTT | TTGGGAACAT | 1260 |

| A AAAAAAA | er ¹ 2. | 1271 |
|--|--------------------|------|
| <210> 30 <211> 1109 <212> DNA <213> Haematococcus pluvialis | | |
| <400> 30 | | |
| TGGAACCTGG CCCGGCGGCA GTCCGATGCC GCGATGCTTC GTTCGTTGCT | | 60 |
| ACGCATATCC CGCGCGTGAA CTCCGCCCAG CAGCCCAGCT GTGCACACGC | | 120 |
| TTTAAGCTCA GGAGCATGCA GCTGCTTGCC GAGGACCGCA CAGACCACAT | GAGGGGTGCA | 180 |
| AGCACCTGGG CAGGCGGCA GTCGCAGGAT GAGCTGATGC TGAAGGACGA | GTGCATCTTA | 240 |
| GTGGATGCTG ACGACAACAT CACAGGCCAT GCCAGCAAGC TGGAGTGCCA | CAAATTCCTA | 300 |
| CCACATCAGC CTGCAGGCCT GCTGCACCGG GCCTTCTCTG TGTTCCTGTT | TGACGACCAG | 360 |
| GGGCGACTGC TGCTGCAACA GCGTGCACGC TCAAAAATCA CCTTCCCAAG | TGTGTGGACG | 420 |
| AACACCTGCT GCAGCCACCC TCTACATGGG CAGACCCCAG ATGAGGTGGA | CCAACTAAGC | 480 |
| CAGGTGGCCG ACGGCACAGT ACCTGGCGCA AAAGCTGCTG CCATCCGCAA | GTTGGAGCAC | 540 |
| GAGCTGGGGA TACCAGCGCA CCAGCTGCCG GCAAGCGCGT TTCGCTTCCT | CACGCGTTTG | 600 |
| CACTACTGTG CCGCGGACGT GCAGCCGGCT GCGACACAAT CAGCGCTCTG | GGGCGAGCAC | 660. |
| GAGATGGACT ACATCTTATT CATCCGGGCC AACGTCACCT TGGCGCCCAA | CCCTGACGAG | 720 |
| GTGGACGAAG TCAGGTACGT GACGCAAGAG GAGCTGCGGC AGATGATGCA | GCCGGACAAC | 780 |
| GGGTTGCAAT GGTCGCCGTG GTTTCGCATC ATCGCCGCGC GCTTCCTTGA | GCGTTGGTGG | 840 |
| GCTGACCTGG ACGCGGCCCT AAACACTGAC AAACACGAGG ATTGGGGAAC | GGTGCATCAC | 900 |
| ATCAACGAAG CGTGAAGGCA GAAGCTGCAG GATGTGAAGA CACGTCATGG | GGTGGAATTG | 960 |
| CGTACTTGGC AGCTTCGTAT CTCCTTTTTC TGAGACTGAA CCTGCAGAGC | TAGAGTCAAT | 1020 |
| GGTGCATCAT ATTCATCGTC TCTCTTTTGT TTTAGACTAA TCTGTAGCTA | GAGTCACTGA | 1080 |
| TGAATCCTTT ACAACTTTCA AAAAAAAAA | | 1109 |
| <210> 31 <211> 985 <212> DNA <213> Lactuca sativa | | |
| <400> 31 TGCCAAAATG TTGAAATTTC CCCCTTTTAA AACCATTGCT ACCATGATCT | CERCE CO. T. | |
| | | 60 |
| TTCTTCCTTC TTGCTGCCTC GGAAATCTTC TTTCCCTCCA ATGCCGTCTC | | 120 |
| TAGTGTTTTC CTCCACCCTC TTTCGTCTGC CGCTATGGGC GATTCCAGCA | | 180 |
| CCAGCGACGT CTCATGTTCG ATGACGAATG CATTTTGGTG GATGAGAATG | | 240 |
| TGGCCATGAT ACTAAATACA ATTGTCATTT GATGGAGAAG ATTGAAAAGG | GAAATATGCT | 300 |

| WO 99/63055 | | |
|-------------|--|--|



| | | | | | • | |
|--|----------------------|------------|-----------------|------------|------------|-----|
| ACACAGAGCA | TTCAGTGTGT | TCTTGTTCAA | CTCGAAATAT | GAATTACTCC | TTCAGCAACG | 360 |
| TTCTGCAACC | AAGGTGACTT | TCCCTTTGGT | ATGGACAAAC | ACGTGTTGCA | GCCATCCACT | 420 |
| ATACAGGGAG | AGTGAGCTTA | TTGACGAAAA | CGCCCTTGGG | GTGAGGAATG | CTGCACAGAG | 480 |
| GAAGCTCCTG | GATGAACTCG | GCATCCCTGG | AGCAGATGTT | CCGGTTGATG | AGTTCACTCC | 540 |
| ATTGGGTCGC | ATTCTATACA | AGGCCGCATC | GGATGGAAAG | TGGGGAGAAC | ATGAACTTGA | 600 |
| TTACCTGCTG | TTTATGGTAC | GTGATGTTGG | TTTGGATCCG | AACCCAGATG | AAGTGAAAGA | 660 |
| TGTAAAATAT | GTGAACCGGG | AAGAGCTGAA | GGAATTGGTA | AGGAAGGCGG | ATGCTGGTGA | 720 |
| AGAGGGTGTG | AAGCTGTCCC | CGTGGTTCAA | ATTGATTGTC | GATAATTTCT | TGTTTCAGTG | 780 |
| GTGGGATCGA | CTCCATAAGG | GAACCCTAAC | CGAAGCTATT | GATATGAAAA | CAATCCACAA | 840 |
| ACTCACATAA | AAACACTACA | CTAGTAGGAG | AGAGGATTAT | ATGAGATATT | TGTTATATGT | 900 |
| GAAATTGAAA | TTCAGATGAA | TGCTTGTATT | TATTTCTATT | TGGACAAACT | TCAACTTCTT | 960 |
| TTTGCTACCT | TATCAGAAAA | AAAAA | | | | 985 |
| <210> 32 <211> 988 <212> DNA <213> Lact | uca sativa | | | | | |
| <400> 32 | 7 7 7 7 M CM CM CM M | | CMC N N M C M C | | | |
| | AAAATCTCTT | | | | | 60 |
| | CACTTTCACC | | | | | 120 |
| | ACGTCTCATG | | | | | 180 |
| | TGATACCAAA | | | | | 240 |
| | AGCATTCAGT | | | | | 300 |
| | AACCAAGGTG | | | | | 360 |
| | AGAATCGGAG | | | | | 420 |
| | | | | | GATGAGTTCA | 480 |
| | | | | | GAACATGAAG | 540 |
| | | | | | GATGAGGTGG | 600 |
| | | | | | GCGGATGCGG | 660 |
| | | | | | TTCTTGTTCA | 720 |
| | TCATGTCCAA | | | | | 780 |
| | | | | | ATAATTTGTG | 840 |
| | | | | | TGGGAGTGTT | 900 |
| TATTGTGTAC | TTGTAACGTA | GGCCCTTTGG | TTACGCTTTA | AGAGTTTAAT | AAAGAACCAC | 960 |
| | | | | | | |

1440

1500

1560

1620



ССТТААТТТА ААААААААА ААААААА 988 <210> 33 <211> 1874 <212> DNA <213> Chlamydomonas reinhardtii GGCACGAGCT CGAGTTTGTT TTACCATGAC ATCGGGAATT TGGAAGCTTG AACTACCTCA 60 ATTACTCAAG TAACTCGCGG CAACACATTT CGCGCGCCAT CGCTGTTTTC TCTGCTCCAG 120 CTACCGAGCA GCATTGCTTT AGATCGCTTT GATGTCATAA ACTCCCACTT ATATGAGATC 180 CAGTTTCATC GAGCCCAAGC CCAGAGCGCA ACCTGTCTTA AGCCGCGGCA GGGCGTCCAT 240 GCGCCTCGCG CAAAGCCGTG CTCTCGTTGC GCGTGTCAGC TCCGCCCTGT GGCCGGGAGC 300 AGGACTTTCA CAGGCTCAAA GCGTTGCGGT GCGAATGGCG AGTTCGTCAA CCTGGGAAGG 360 CACGGGCCTG AGCCAGGATG ACTTCATGCA GCGGGACGAG TGCTTGGTGG TGGACGAGCA 420 GGACCGGCTG CTAGGCACCG CCAACAAGTA CGACTGCCAC CGCTTCGAGG CGGCCAAGGG 480 CCAGCCCTGC GGCCGCCTGC ACCGCGCCTT CTCCGTGTTC CTGTTCAGCC CCGACGGCCG 540 ACTGCTGCTG CAGCAGCGC CAGCCAGCAA GGTGACGTTC CCGGGTGTGT GGACCAACAC 600 CTGCTGCTCG CACCCGCTGG CGGGCCAGGC GCCGGACGAG GTGGACCTGC CGGCGGCGGT 660 AGCCTCGGGC CAGGTGCCGG GCATCAAGGC GGCGGCGGTG CGCAAGCTGC AGCACGAGCT 720 GGGGATACCG CCGGAGCAGG TTCCCGCCTC CTCCTTCTCC TTCCTCACGC GTCTGCACTA 780 CTGCGCCGCC GACACCGCCA CGCACGGCCC GGCGGCGGAG TGGGGCGAGC ACGAGGTGGA 840 CTACGTGCTG TTCGTGCGGC CGCAGCAGCC CGTCAGCCTG CAGCCCAACC CAGACGAGGT 900 GGACGCCACG CGCTACGTGA CGCTGCCGGA GCTTCAGTCC ATGATGGCGG ACCCCGGCCT 960 CAGCTGGAGC CCCTGGTTCC GCATCCTGGC CACACAGCCC GCCTTCCTGC CCGCCTGGTG 1020 GGGCGACCTG AAGCGGCGCT GGCGCCCGGG CGGCAGCCGA CTGTCGGACT GGGGCACCAT 1080 CCACCGCGTC ATGTGAAGAA AAAGGGGAAG CAGGGGGGGG AGCGGGGGAT GAATGGGAAT 1140 GTGAATGCGA TTGTGATGCG GCGTGGGATG AGGTCTGAAG ACAGGGGGAA AATCGGGGGG 1200 CGGGCGTGAG CGTGTGTGTA CGTGAGCGAC AAAGCCGGGA GGCGGACCGC GCGATGGGTA 1260 CATGTGTGTG CGGAGGGTCG GTGGGTCGGT CGGTTGCGCG GCATAGCGTG TTGTGTGTGT 1320 GCGGCTGCAG GGGTATGTGG GCACCCGGGC ACGGAGGAGA AGGCACACGC AGGTGGCGCG 1380

GAGGTGTGTC AGGGGCCCATG GGCGGGCCTC ACTCCTGGTC GTGCCCAGTG GTCTCGTGGG

CAGAGTGGCA GGGGCTGCAC CCATATGAGC GGCGCACTGC CGCGCTGGGC TAAGTCCTTA

TCACTTGGTG AGGTGGGCG AGGTGGCTGT GGGCGGCGGG CGCAGTGGCA GAAGGACACG

GTGTGTGAGC GGTGGAGCTC TGGCCGTGCC GGCCGTGAGG GGCGGATAGC GATATGACGT

| WO 99/63055 | | PCT/US99/12121 |
|-------------|--|----------------|
| | | |

| TGTGCTTGGC CGCTGTAATG | CGGGAGAATG | TGCAGGCCGC | GAGAAGCGGG | CGGTGGCAGG | 1680 |
|--|------------|------------|------------|------------|------|
| AGGCCGCAGG CTGCAGCACC | CGTTĠGGGAG | GTGCCACCTG | CAGGCGCGGC | GCCGGGCGGG | 1740 |
| CCTGAGTAAT GGGCGCCTGA | GTAGTGGCGG | CCACAGGAGG | CGCAGGAGGC | AGCAGCAGGA | 1800 |
| GGACGAGCTG GAGGGACCCG | TTGGCAACCC | AAGGTTGCGC | GTGTAACATA | GTGGCCATAC | 1860 |
| AAAAAAAAAA | | | | | 1874 |
| <210> 34 <211> 954 <212> DNA <213> Tagetes erecta | | | | • | · |
| <400> 34 CCAAAAACAA CTCAAATCTC | СТСССТСССТ | CTTACTCCC | CARCCCRCAC | Ch Cmccccc | |
| TGGATGCTGT TCAGCGACGT | | | | | 60 |
| | | | | | 120 |
| ACAATGTGGT GGGACATGAT | | | | | 180 |
| GTAAAATGCT GCACAGAGCA | | | | | 240 |
| TTCAGCAACG GTCTGCAACC | | | | | 300 |
| GCCATCCACT CTACAGAGAA | TCCGAGCTTG | TTCCCGAAAA | CGCCCTTGGA | GTAAGAAATG | 360 |
| CTGCACAGAG GAAGCTGTTG | GATGAACTCG | GTATCCCTGC | TGAAGATGTT | CCCGTTGATC | 420 |
| AGTTTACTCC TTTAGGTCGC | ATGCTCTACA | AGGCTCCATC | TGATGGAAAG | TGGGGAGAAC | 480 |
| ATGAACTTGA CTACCTACTT | TTCATAGTGA | GAGACGTTGC | TGTAAACCCG | AACCCAGATG | 540 |
| AAGTGGCGGA TATCAAATAT | GTGACCAGAA | GAGTTAAAGG | AGCTGCTAAG | GAAAGCAGAT | 600 |
| GCGGGGGAGG AGGGTTTGAA | GCTGTCTCCA | TGGTTCAGGT | TAGTGGTTGA | TAACTTCTTG | 660 |
| TTCAAGTGGT GGGATCATGT | GCAAAAGGGT | ACACTCACTG | AAGCAATTGA | TATGAAAACC | 720 |
| ATACACAAGC TGATATAGAA | ACACACCCTC | AACCGAAAAG | TTCAAGCCTA | ATAATTCGGG | 780 |
| TTGGGTCGGG TCTACCATCA | ATTGTTTTTT | TCTTTTAAGA | AGTTTTAATC | TCTATTTGAG | 840 |
| CATGTTGATT CTTGTCTTTT | GTGTGTAAGA | TTTTGGGTTT | CGTTTCAGTT | GTAATAATGA | 900 |
| ACCATTGATG GTTTGCAATT | TCAAGTTCCT | ATCGACATGT | AGTGATCTAA | AAAA | 954 |
| <210> 35 <211> 1031 <212> DNA <213> Oryza sativa | | | | | |
| <400> 35 | 20005 | | | | |
| CCTCCCTTTG CCTCGCGCAG | | | | | 60 |
| GCCGCCGCCG TGGAGGACGC | | | | | 120 |
| GAATGCATTT TGGTGGATGA | ACAAGACAAT | GTTGTTGGCC | ATGAATCAAA | ATATAACTGC | 180 |
| CATCTGATGG AAAAAATCGA | ATCTGAAAAT | CTACTTCATA | GGGCTTTCAG | TGTATTCCTG | 240 |

WO 99/63055



PCT/US99/12121

| | | | | | · | |
|------------|------------|------------|------------|------------|------------|------|
| TTCAACTCAA | AATATGAACT | CCTACTCCAG | CAACGATCTG | CAACAAAGGT | TACATTTCCT | 300 |
| CTAGTTTGGA | CCAACACTTG | CTGCAGCCAT | CCTCTGTACC | GTGAGTCTGA | GCTTATACAG | 360 |
| GAAAACTACC | TTGGTGTTAG | AAATGCTGCT | CAGAGGAAGC | TCTTGGATGA | GCTGGGCATC | 420 |
| CCAGCTGAAG | ATGTGCCAGT | TGACCAATTC | ACCCCTCTTG | GTCGGATGCT | TTACAAGGCC | 480 |
| CCATCTGATG | GAAAATGGGG | TGAACACGAG | CTTGACTACC | TGCTGTTCAT | CGTCCGCGAC | 540 |
| GTGAAGGTAG | TCCCGAACCC | GGACGAAGTG | GCCGATGTGA | AATACGTGAG | CCGTGAGCAG | 600 |
| CTGAAGGAGC | TCATCCGCAA | AGCGGACGCC | GGAGAGGAAG | GCCTGAAGCT | GTCTCCCTGG | 660 |
| TTCCGGCTGG | TTGTTGACAA | CTTCCTCATG | GGCTGGTGGG | ATCACGTCGA | GAAAGGCACC | 720 |
| CTCAACGAGG | CCGTGGACAT | GGAGACCATC | CACAAGCTGA | AGTAAGGACT | GCGATGTTGT | 780 |
| GGCTGGAAAG | AATGATCCTG | AAGACTCTGT | тсттстсстс | CTGCATATTA | CTCTTACCAG | 840 |
| GGAAGTTGCA | GAAGTCAGAA | GAAGCTTTTG | TATGTTTCTG | GGTTTGGAGC | TTGGAAGTGT | 900 |
| TGGGCTCTGC | TGACTGAGAG | ATTCCCTTAT | AGAGTGTCTA | TGTTAATTTA | GCAAACTTCT | 960 |
| ATATTATACA | TGATTAGTTA | ATTGTTCGGT | GTCTGAATAA | AGAACAATAG | CATGTTCCAT | 1020 |
| GTTTATTTGC | Т | | | | | 1031 |
| | | | | • | | |

<210> 36

<211> 232

<212> PRT

<213> Tagetes erecta

<400> 36

Met Gly Asp Asp Ser Gly Met Asp Ala Val Gln Arg Arg Leu Met Phe 1 5 10 15

Asp Asp Glu Cys Ile Leu Val Asp Glu Cys Asp Asn Val Val Gly His 20 25 30

Asp Thr Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Thr Gly Lys 35 40 45

Met Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu
50 55 60

Leu Leu Cln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro Leu Val
65 70 75 80

Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser Glu Leu 85 90 95

Val Pro Glu Asn Ala Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu 100 105 110

Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp Gln Phe
115 120 125

Thr Pro Leu Gly Arg Met Leu Tyr Lys Ala Pro Ser Asp Gly Lys Trp 130 135 140



- Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Ile Val Arg Asp Val Ala 145 150 155 160
- Val Asn Pro Asn Pro Asp Glu Val Ala Asp Ile Lys Tyr Val Ser His 165 170 175
- Glu Glu Leu Lys Glu Leu Leu Arg Lys Ala Asp Ala Gly Glu Glu Gly 180 185 190
- Leu Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe Leu Phe 195 200 205
- Lys Trp Trp Asp His Val Gln Lys Gly Thr Leu Thr Glu Ala Ile Asp 210 215 220
- Met Lys Thr Ile His Lys Leu Ile 225 230
- <210> 37
- <211> 280
- <212> PRT
- <213> Lactuca Sativa
- <400> 37
- Met Leu Lys Phe Pro Pro Phe Lys Thr Ile Ala Thr Met Ile Ser Ser 1 5 10 15
- Pro Tyr Ser Ser Phe Leu Leu Pro Arg Lys Ser Ser Phe Pro Pro Met 20 25 30
- Pro Ser Leu Ala Ala Ser Val Phe Leu His Pro Leu Ser Ser Ala 35 40 45
- Ala Met Gly Asp Ser Ser Met Asp Ala Val Gln Arg Arg Leu Met Phe 50 55 60
- Asp Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Lys Val Val Gly His 65 70 75 80
- Asp Thr Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Lys Gly Asn 85 90 95
- Met Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu 100 105 110
- Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro Leu Val 115 120 125
- Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser Glu Leu 130 135 140
- Ile Asp Glu Asn Ala Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu 145 150 155 160
- Leu Asp Glu Leu Gly Ile Pro Gly Ala Asp Val Pro Val Asp Glu Phe
 165 170 175
- Thr Pro Leu Gly Arg Ile Leu Tyr Lys Ala Ala Ser Asp Gly Lys Trp 180 185 190
- Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Met Val Arg Asp Val Gly

 $\sigma^{i_{1}}$

195 200 205

Leu Asp Pro Asn Pro Asp Glu Val Lys Asp Val Lys Tyr Val Asn Arg

Glu Glu Leu Lys Glu Leu Val Arg Lys Ala Asp Ala Gly Glu Gly

Val Lys Leu Ser Pro Trp Phe Lys Leu Ile Val Asp Asn Phe Leu Phe

Gln Trp Trp Asp Arg Leu His Lys Gly Thr Leu Thr Glu Ala Ile Asp 260 265

Met Lys Thr Ile His Lys Leu Thr

<210> 38

<211> 229

<212> PRT

<213> Lactuca Sativa

<400> 38

Met Gly Asp Asp Ser Gly Met Asp Ala Val Gln Arg Arg Leu Met Phe

Asp Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Asn Val Leu Gly His

Asp Thr Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Lys Asp Asn

Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu

Leu Leu Leu Gln Gln Arg Ser Glu Thr Lys Val Thr Phe Pro Leu Val

Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser Glu Leu

Ile Pro Glu Asn Ala Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu

Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp Glu Phe

Thr Thr Leu Gly Arg Met Leu Tyr Lys Ala Pro Ser Asp Gly Lys Trp

Gly Glu His Glu Val Asp Tyr Leu Leu Phe Leu Val Arg Asp Val Ala 150

Val Asn Pro Asn Pro Asp Glu Val Ala Asp Ile Arg Tyr Val Asn Gln

Glu Glu Leu Lys Glu Leu Leu Arg Lys Ala Asp Ala Gly Glu Glu Gly

Leu Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe Leu Phe 195



Lys Trp Trp Asp His Val Gln Lys Gly Thr Leu Asn Glu Ala Ile Asp 210 225 220

Met Lys Thr Ile His

<210> 39

<211> 295

<212> PRT

<213> Adonis Palaestina

<400> 39

BNSDOCID: <WO_

9963055A1 IA>

Met Ser Ser Ile Arg Ile Asn Pro Leu Tyr Ser Ile Phe Ser Thr Thr 1 5 10 15

Thr Lys Thr Leu Ser Ala Ser Cys Ser Ser Pro Ala Val His Leu Gln 20 25 30

Gln Arg Cys Arg Thr Leu Ser Ile Ser Ser Ser Ile Thr Asn Ser Pro 35 40 45

Arg Arg Gly Leu Asn Arg Leu Phe Ala Ser Thr Ser Thr Met Gly Glu 50 60

Val Ala Asp Ala Gly Met Asp Ala Val Gln Lys Arg Leu Met Phe Asp 65 70 75 80

Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Lys Val Val Gly Tyr Asp 85 90 95

Ser Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala Glu Asn Leu 100 105 110

Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys Tyr Glu Leu 115 120 125

Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro Leu Val Trp 130 135 140

Thr Asn Thr Cys Cys Ser His Pro Leu Phe Arg Asp Ser Glu Leu Ile 145 150 155 160

Glu Glu Asn Phe Leu Gly Val Arg Asn Ala Ala Gln Arg Lys Leu Leu 165 170 175

Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp Glu Phe Thr 180 185 190

Pro Leu Gly Arg Ile Leu Tyr Lys Ala Pro Ser Asp Gly Lys Trp Gly 195 200 205

Glu His Glu Leu Asp Tyr Leu Leu Phe Ile Val Arg Asp Val Lys Tyr 210 215 220

Asp Pro Asn Pro Asp Glu Val Ala Asp Ala Lys Tyr Val Asn Arg Glu 225 230 235 240

Glu Leu Lys Glu Ile Leu Arg Lys Ala Asp Ala Gly Glu Glu Gly Ile 245 250 255

Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe Leu Phe Lys

260

265

270

Trp Trp Asp His Val Glu Glu Gly Lys Ile Lys Asp Val Ala Asp Met 275 280 285

Lys Thr lle His Lys Leu Thr 290 295

<210> 40

<211> 234

<212> PRT

<213> Adonis Palaestina

<400> 40

Met Gly Glu Val Thr Asp Ala Gly Met Asp Ala Val Gln Lys Arg Leu $1 \hspace{1cm} 5 \hspace{1cm} 10 \hspace{1cm} 15$

Met Phe Asp Asp Glu Cys Ile Leu Val Asp Glu Asn Asp Lys Val Val 20 25 30

Gly His Asp Ser Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala 35 40 45

Glu Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys
50 55 60

Tyr Glu Leu Leu Gln Gln Arg Ser Ala Thr Lys Val Thr Phe Pro 65 70 75 80

Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu Phe Arg Asp Ser 85 90 95

Glu Leu Ile Glu Glu Asn Tyr Leu Gly Val Arg Asn Ala Ala Gln Arg 100 105 110

Lys Leu Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp Val Pro Val Asp 115 120 125

Glu Phe Thr Pro Leu Gly Arg Ile Leu Tyr Lys Ala Pro Ser Asp Gly 130 135

Lys Trp Gly Glu His Glu Leu Asp Tyr Leu Leu Phe Ile Val Arg Asp 145 150 155 160

Val Lys Tyr Asp Pro Asn Pro Asp Glu Val Ala Asp Ala Lys Tyr Val 165 170 175

Asn Arg Glu Glu Leu Arg Glu Ile Leu Arg Lys Ala Asp Ala Gly Glu 180 185 190

Glu Gly Leu Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe 195 200 205

Leu Phe Lys Trp Trp Asp His Val Glu Gln Gly Thr Ile Lys Glu Val 210 215 220

Ala Asp Met Lys Thr Ile His Lys Leu Thr 225 230

<210> 41 <211> 238



<212> PRT <213> Oryza Sativa

Gln Lys Arg Leu Met Phe Asp Asp Glu Cys Ile Leu Val Asp Glu Gln 20 25 30

Asp Asn Val Val Gly His Glu Ser Lys Tyr Asn Cys His Leu Met Glu 35 40 45

Lys Ile Glu Ser Glu Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu 50 55 60

Phe Asn Ser Lys Tyr Glu Leu Leu Gln Gln Arg Ser Ala Thr Lys 65 70 75 80

Val Thr Phe Pro Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu 85 90 95

Tyr Arg Glu Ser Glu Leu Ile Gln Glu Asn Tyr Leu Gly Val Arg Asn 100 105 110

Ala Ala Gln Arg Lys Leu Leu Asp Glu Leu Gly Ile Pro Ala Glu Asp 115 120 125

Val Pro Val Asp Gln Phe Thr Pro Leu Gly Arg Met Leu Tyr Lys Ala 130 135 140

Pro Ser Asp Gly Lys Trp Gly Glu His Glu Leu Asp Tyr Leu Leu Phe 145 150 155 160

Ile Val Arg Asp Val Lys Val Val Pro Asn Pro Asp Glu Val Ala Asp 165 170 175

Val Lys Tyr Val Ser Arg Glu Gln Leu Lys Glu Leu Ile Arg Lys Ala 180 - 185 190

Asp Ala Gly Glu Glu Gly Leu Lys Leu Ser Pro Trp Phe Arg Leu Val 195 200 205

Val Asp Asn Phe Leu Met Gly Trp Trp Asp His Val Glu Lys Gly Thr 210 215 220

Leu Asn Glu Ala Val Asp Met Glu Thr Ile His Lys Leu Lys 225 230 235

<210> 42

<211> 233

<212> PRT

<213> Arabidopsis thaliana

<400> 42

Met Thr Asp Ser Asn Asp Ala Gly Met Asp Ala Val Gln Arg Arg Leu
1 5 10 15

Met Phe Glu Asp Glu Cys Ile Leu Val Asp Glu Asn Asn Arg Val Val 20 25 30



Gly His Asp Thr Lys Tyr Asn Cys His Leu Met Glu Lys Ile Glu Ala 35 40 45

Glu Asn Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asn Ser Lys 50 55 60

Tyr Glu Leu Leu Gln Gln Arg Ser Lys Thr Lys Val Thr Phe Pro 65 70 75 80

Leu Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu Tyr Arg Glu Ser 85 90 95

Glu Leu Ile Glu Glu Asn Val Leu Gly Val Arg Asn Ala Ala Gln Arg 100 105 110

Lys Leu Phe Asp Glu Leu Gly Ile Val Ala Glu Asp Val Pro Val Asp 115 120 125

Glu Phe Thr Pro Leu Gly Arg Met Leu Tyr Lys Ala Pro Ser Asp Gly 130 135 140

Lys Trp Gly Glu His Glu Val Asp Tyr Leu Leu Phe Ile Val Asp 145 155 160

Val Lys Leu Gln Pro Asn Pro Asp Glu Val Ala Glu Ile Lys Tyr Val 165 170 175

Ser Arg Glu Glu Leu Lys Glu Leu Val Lys Lys Ala Asp Ala Gly Asp 180 185 190

Glu Ala Val Lys Leu Ser Pro Trp Phe Arg Leu Val Val Asp Asn Phe 195 200 205

Leu Met Lys Trp Trp Asp His Val Glu Lys Gly Thr Ile Thr Glu Ala 210 215 220

Ala Asp Met Lys Thr Ile His Lys Leu 225 230

<210> 43

<211> 293

<212> PRT

<213> Haematococcus pluvialis

<400> 43

Met Leu Arg Ser Leu Leu Arg Gly Leu Thr His Ile Pro Arg Val Asn 1 5 10 15

Ser Ala Gln Gln Pro Ser Cys Ala His Ala Arg Leu Gln Phe Lys Leu 20 25 30

Arg Ser Met Gln Leu Leu Ser Glu Asp Arg Thr Asp His Met Arg Gly
35 40 45

Ala Ser Thr Trp Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys
50 55 60

Asp Glu Cys Ile Leu Val Asp Val Glu Asp Asn Ile Thr Gly His Ala 65 70 75 80

Ser Lys Leu Glu Cys His Lys Phe Leu Pro His Gln Pro Ala Gly Leu

. 85

Leu His Arg Ala Phe Ser Val Phe Leu Phe Asp Asp Gln Gly Arg Leu 100 105 110

Leu Leu Gln Gln Arg Ala Arg Ser Lys Ile Thr Phe Pro Ser Val Trp 115 120 125

Thr Asn Thr Cys Cys Ser His Pro Leu His Gly Gln Thr Pro Asp Glu 130 135 140

Val Asp Gln Leu Ser Gln Val Ala Asp Gly Thr Val Pro Gly Ala Lys 145 150 155 160

Ala Ala Ala Ile Arg Lys Leu Glu His Glu Leu Gly Ile Pro Ala His 165 170 175

Gln Leu Pro Ala Ser Ala Phe Arg Phe Leu Thr Arg Leu His Tyr Cys 180 185 190

Ala Ala Asp Val Gln Pro Ala Ala Thr Gln Ser Ala Leu Trp Gly Glu 195 200 205

His Glu Met Asp Tyr Ile Leu Phe Ile Arg Ala Asn Val Thr Leu Ala 210 215 220

Pro Asn Pro Asp Glu Val Asp Glu Val Arg Tyr Val Thr Gln Glu Glu 225 230 235 240

Leu Arg Gln Met Met Gln Pro Asp Asn Gly Leu Gln Trp Ser Pro Trp 245 250 255

Phe Arg Ile Ile Ala Ala Arg Phe Leu Glu Arg Trp Trp Ala Asp Leu 260 265 270

Asp Ala Ala Leu Asn Thr Asp Lys His Glu Asp Trp Gly Thr Val His

His Ile Asn Glu Ala 290

<210> 44

<211> 304

<212> PRT

<213> Haematococcus pluvialis

<400> 44

Met Leu Arg Ser Leu Leu Arg Gly Leu Thr His Ile Pro Arg Val Asn 1 5 10 15

Ser Ala Gln Gln Pro Ser Cys Ala His Ala Arg Leu Gln Phe Lys Leu 20 25 30

Arg Ser Met Gln Met Thr Leu Met Gln Pro Ser Ile Ser Ala Asn Leu 35 40 45

Ser Arg Ala Glu Asp Arg Thr Asp His Met Arg Gly Ala Ser Thr Trp 50 55 60

Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys Asp Glu Cys Ile 65 70 75 80

WO 99/63055



Leu Val Asp Val Glu Asp Asn Ile Thr Gly His Ala Ser Lys Leu Glu 85 90 95

Cys His Lys Phe Leu Pro His Pro Ala Gly Leu Leu His Arg Ala Phe 100 105 110

Ser Val Phe Leu Phe Asp Asp Gln Gly Arg Leu Leu Gln Gln Arg 115 120 125

Ala Arg Ser Lys Ile Thr Phe Pro Ser Val Trp Thr Asn Thr Cys Cys 130 135 140

Ser His Pro Leu His Gly Gln Thr Pro Asp Glu Val Asp Gln Leu Ser 145 150 155 160

Gln Val Ala Asp Gly Thr Val Pro Gly Ala Lys Ala Ala Ala Ile Arg 165 170 175

Lys Leu Glu His Glu Leu Gly Ile Pro Ala His Gln Leu Pro Ala Ser 180 185 190

Ala Phe Arg Phe Leu Thr Arg Leu His Tyr Cys Ala Ala Asp Val Gln
195 200 205

Pro Ala Ala Thr Gln Ser Ala Leu Trp Gly Glu His Glu Met Asp Tyr 210 215 220

Ile Leu Phe Ile Arg Ala Asn Val Thr Leu Ala Pro Asn Pro Asp Glu 225 230 235 240

Val Asp Glu Val Arg Tyr Val Thr Gln Glu Glu Leu Arg Gln Met Met
245 250 255

Gln Pro Asp Asn Gly Leu Gln Trp Ser Pro Trp Phe Arg Ile Ile Ala 260 265 270

Ala Arg Phe Leu Glu Arg Trp Trp Ala Asp Leu Asp Ala Ala Leu Asn 275 280 285

Thr Asp Lys His Glu Asp Trp Gly Thr Val His His Ile Asn Glu Ala 290 295 300

<210> 45

<211> 307

<212> PRT

<213> Chlamydomonas reinhardtii

<400> 45

Met Arg Ser Ser Phe Ile Glu Pro Lys Pro Arg Ala Gln Pro Val Leu 1 15

Ser Arg Gly Arg Ala Ser Met Arg Leu Ala Gln Ser Arg Ala Leu Val 20 25 30

Ala Arg Val Ser Ser Ala Leu Trp Pro Gly Ala Gly Leu Ser Gln Ala 35 40 45

Gln Ser Val Ala Val Arg Met Ala Ser Ser Ser Thr Trp Glu Gly Thr
50 55 60

Gly Leu Ser Gln Asp Asp Phe Met Gln Arg Asp Glu Cys Leu Val Val

70 75 Asp Glu Gln Asp Arg Leu Leu Gly Thr Ala Asn Lys Tyr Asp Cys His Arg Phe Glu Ala Ala Lys Gly Gln Pro Cys Gly Arg Leu His Arg Ala 105 Phe Ser Val Phe Leu Phe Ser Pro Asp Gly Arg Leu Leu Gln Gln Arg Ala Ala Ser Lys Val Thr Phe Pro Gly Val Trp Thr Asn Thr Cys 135 Cys Ser His Pro Leu Ala Gly Gln Ala Pro Asp Glu Val Asp Leu Pro Ala Ala Val Ala Ser Gly Gln Val Pro Gly Ile Lys Ala Ala Ala Val Arg Lys Leu Gln His Glu Leu Gly Ile Pro Pro Glu Gln Val Pro Ala Ser Ser Phe Ser Phe Leu Thr Arg Leu His Tyr Cys Ala Ala Asp Thr 200 Ala Thr His Gly Pro Ala Ala Glu Trp Gly Glu His Glu Val Asp Tyr Val Leu Phe Val Arg Pro Gln Gln Pro Val Ser Leu Gln Pro Asn Pro 235 Asp Glu Val Asp Ala Thr Arg Tyr Val Thr Leu Pro Glu Leu Gln Ser 250 Met Met Ala Asp Pro Gly Leu Ser Trp Ser Pro Trp Phe Arg Ile Leu Ala Thr Gln Pro Ala Phe Leu Pro Ala Trp Trp Gly Asp Leu Lys Arg Arg Trp Arg Pro Gly Gly Ser Arg Leu Ser Asp Trp Gly Thr Ile His 300 295 Arg Val Met 305 <210> 46 <211> 1848 <212> DNA <213> Adonis palaestina <400> 46 GAGAGAAAA GAGTGTTATA TTAATGTTAC TGTCGCATTC TTGCAACACA TATTCAGACT 60 CCATTTTCTT GTTTTCTCTT CAAAACAACA AACTAATGTG ACGGAGTATC TAGCTATGGA 120 ACTACTTGGT GTTCGCAACC TCATCTCTTC TTGCCCTGTC TGGACTTTTG GAACAAGAAA 180 240

WO 99/63055



| TTTTCAAGTG | AGGGCTGATG | GTGGAAGCGG | GAGTAGAACT | TCTGTTGCTT | ATAAAGAGGG | 300 |
|------------|-------------|--------------|--------------|-------------|--------------|------|
| TTTTGTGGAC | GAGGAGGATT | TTATCAAAGC | TGGTGGTTCT | GAGCTTTTGT | TTGTCCAAAT | 360 |
| GCAGCAAACA | AAGTCTATGG | AGAAACAGGC | CAAGCTCGCC | GATAAGTTGC | CACCAATACC | 420 |
| TTTCGGAGAA | TCTGTGATGG | ACTTGGTTGT | AATAGGTTGT | GGACCTGCTG | GTCTTTCACT | 480 |
| GGCTGCAGAA | GCTGCTAAGC | TAGGCTTGAA | AGTTGGCCTT | ATTGGTCCTG | ATCTTCCTTT | 540 |
| TACAAATAAT | TATGGTGTGT | GGGAAGACGA | GTTCAAAGAT | CTTGGACTTG | AACGTTGTAT | 600. |
| CGAGCATGCT | TGGAAGGACA | CCATCGTATA | TCTTGACAAT | GATGCTCCTG | TCCTTATTGG | 660 |
| TCGTGCATAT | GGACGAGTTA | GCCGGCATTT | GCTGCATGAA | GAGTTGCTGA | AAAGGTGTGT | 720 |
| CGAGTCAGGT | GTATCATATC | TGAATTCTAA | AGTGGAAAGG | ATCACTGAAG | CTGGTGATGG | 780 |
| CCATAGTCTT | GTAGTTTGTG | AAAACGACAT | CTTTATCCCT | TGCAGGCTTG | CTACTGTTGC | 840 |
| ATCTGGAGCA | GCTTCAGGGA | AACTTTTGGA | GTATGAAGTA | GGTGGCCCTC | GTGTTTGTGT | 900 |
| CCAAACTGCT | TATGGTGTGG | AGGTTGAGGT | GGAGAACAAT | CCATACGATC | CCAACTTAAT | 960 |
| GGTATTTATG | GACTACAGAG | ACTATATGCA | ACAGAAATTA | CAGTGCTCGG | AAGAAGAATA | 1020 |
| TCCAACATTT | CTCTATGTCA | TGCCCATGTC | GCCAACAAGA | CTTTTTTTTG | AGGAAACCTG | 1080 |
| TTTGGCCTCA | AAAGATGCCA | TGCCTTTCGA | TCTACTGAAG | AGAAAACTAA | TGTCACGATT | 1140 |
| GAAGACTCT | GGTATCCAAC | TTACAAAAAT | TTATGAAGAG | GAATGGTCTT | ATATTCCTGT | 1200 |
| TGGGGGTTCT | TTACCAAAC | A CAGAGCAAAA | GAACCTAGCA | TTTGGTGCTG | CAGCAAGCAT | 1260 |
| GGTGCATCC | GCAACAGGC | r ATTCGGTTGI | ACGATCACTA | TCAGAAGCT | CAAAATATGC | 1320 |
| TTCTGTAAT | GCAAAGATT | r TGAAGCAAGA | A TAACTCTGCA | A TATGTGGTT | CTGGACAAAG | 1380 |
| CAGTGCAGT | A AACATTTCA | A TGCAAGCATO | G GAGCAGTCTT | TGGCCAAAG | S AGCGAAAACG | 1440 |
| TCAAAGAGC | A TTCTTTCTT | T TCGGGTTAGA | A GCTTATTGTO | CAGCTAGAT | A TTGAAGCAAC | 1500 |
| CAGAACGTT | C TTTAGAACC | T TCTTCCGCT | r GCCAACTTGO | ATGTGGTGG | GGTTTCCTTGG | 1560 |
| GTCTTCACT. | A TCATCTTTC | G ATCTTGTAT | T GTTTTCCAT(| G TACATGTÍT | G TTTTGGCCCC | 1620 |
| GAACAGCAT | g AGGATGTCA | C TTGTGAGAC | A TTTGCTTTC | A GATCCTTCT | G GTGCAGTTAT | 1680 |
| GGTTAAAGC | T TACCTCGAA | A GGTAATCTG | T TTTATGAAA | C TATAGTGTC | т саттааатаа | 1740 |
| ATGAGGATC | C TTCGTATAT | G TATATGATC | A TCTCTATGT | А ТАТССТАТА | т тстаатстса | 1800 |
| TAAAGTAAT | C GAAAATTCA | T TGATAGAAA | AAAAAAA A | AAAAAAA A | • | 1848 |

<210> 47

Met Glu Leu Cly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp

<211> 529

<212> PRT

^{. &}lt;213> Adonis palaestina

<400> 47

#2

| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
|------------|------------|--------------|--------------|--------------|--------------|------------|------------|------------|------------|------------|------------|--------------|------------------------|------------|------------|
| Thr | Phe | Gly | Thr 20 | Arg | Asn : | Leu | Ser | Ser 25 | Ser | Lys | Leu | Ala | Tyr 30 | Asn | Ile |
| His. | Arg | Tyr 35 | Gly | Ser | Ser | Cys | Arg 40 | Val | Asp | Phe | Gln | Val 45 | Arg | Ala | Asp |
| Gly | Gly 50 | Ser | Gly | Ser | Arg | Ser 55 | Ser | Val | Ala | Tyr | Lys 60 | Glu | Gly | Phe | Val |
| Asp 65 | Glu | Glu | Asp | Phe | Ile 70 | Lys | Ala | Gly | Gly | Ser 75 | Glu | Leu | Leu | Phe | Val 80 |
| Gln | Met | Gln | Gln | Thr 85 | Lys | Ser | Met | Glu | Lys 90 | Gln | Ala | Lys | Leu | Ala 95 | Asp |
| Lys | Leu | Pro | Pro 100 | Ile | Pro | Phe | Gly | Glu 105 | Ser | Val | Met | Asp | Leu 110 | Val | Val |
| Ile | Gly | Cys 115 | Gly | Pro | Ala | Gly | Leu 120 | Ser | Leu | Ala | Ala | Glu 125 | Ala | Ala | Lys |
| Leu | Gly 130 | Leu | Lys | Val | Gly | Leu 135 | Ile | Gly | Pro | Asp | Leu 140 | Pro | Phe | Thr | Asn |
| Asn 145 | Tyr | Gly | Val | Trp | Glu 150 | Asp | Glu | Phe | Lys | Asp 155 | Leu | Gly | Leu | Glu | Arg 160 |
| Cys | Ile | Glu | His | Ala 165 | Trp | Lys | Asp | Thr | Ile 170 | Val | Tyr | Leu | Asp | Asn 175 | Asp |
| Ala | Pro | Val | Leu 180 | | Gly | Arg | Ala | Tyr 185 | Gly | Arg | Val | Ser | Arg 190 | His | Leu |
| Leu | His | Glu 195 | | Leu | Leu | Lys | Arg 200 | | Val | Glu | Ser | Gly 205 | Val | Ser | Tyr |
| Leu | Asp 210 | | Lys | : Val | Glu | Arg 215 | | Thr | Glu | Ala | Gly 220 | Asp | Gly | His | Ser |
| Leu 225 | | Val | . Cys | s Glu | Asn 230 | | ılle | Phe | e Ile | 235 | Cys | s Arg | g Leu | Ala | Thr 240 |
| Val | Ala | Ser | Gly | / Ala 245 | a Ala | Ser | Gly | , Lys | 250 | ı Let | ı Glu | тул туп | Glu | val 255 | Gly |
| Gly | Pro | Arg | y Val 260 | | s Val | . Glr | n Thr | 265 | а Туз 5 | r Gly | y Vai | l Glı | u Val 270 | l Glu | ı Val |
| Glu | Ası | n Ası 27! | | э Туі | r Asp | Pro | 280 | | u Met | t Vai | l Ph | e Met 28: | t As _f 5 | э Туі | r Ar |
| Asp | ту: 290 | _ | t Gli | n Gli | n Lys | 29 | | n Cy: | s Se | r Gl | u Gl 30 | u Gl: 0 | и Ту: | r Pro | o Th |
| Phe 305 | | и Ту | r Va | l Me | t Pro 310 | | t Se | r Pr | o Th | r Ar 31 | g Le 5 | u Ph | e Ph | e Gl | u G1 32 |
| Th | c Cy | s Le | u Al | a Se 32 | r Ly: 5 | s As | p Al | a Me | t Pr 33 | o Ph O | e As | p Le | u Le | u Ly 33 | s Ar 5 |



Lys Leu Met Ser Arg Leu Lys Thr Leu Gly Ile Gln Val Thr Lys Val 340 345 350

Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn 355 360 365

Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val His $370 \hspace{1.5cm} 375 \hspace{1.5cm} 380$

Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys 385 390 395 400

Tyr Ala Ser Val Ile Ala Lys Ile Leu Lys Gln Asp Asn Ser Ala Tyr 405 410 415

Val Val Ser Gly Gln Ser Ser Ala Val Asn Ile Ser Met Gln Ala Trp 420 425 430

Ser Ser Leu Trp Pro Lys Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu 435 440 445

Phe Gly Leu Glu Leu Ile Val Gln Leu Asp Ile Glu Ala Thr Arg Thr 450 455 460

Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp Trp Gly Phe 465 470 475 480

Leu Gly Ser Ser Leu Ser Ser Phe Asp Leu Val Leu Phe Ser Met Tyr 485 490 495

Met Phe Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His 500 505 510

Leu Leu Ser Asp Pro Ser Gly Ala Val Met Val Arg Ala Tyr Leu Glu 515 520 525

Ara

<210> 48

<211> 378

<212> PRT

<213> Potato

<400> 48

Asp Glu Phe Lys Asp Leu Gly Leu Gln Ala Cys Ile Glu His Val Trp
1 5 10 15

Arg Asp Thr Ile Val Tyr Leu Asp Asp Asp Pro Ile Leu Ile Gly
20 25 30

Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu His Glu Glu Leu Leu 35 40 45

Lys Arg Cys Val Glu Ala Gly Val Leu Tyr Leu Asn Ser Lys Val Asp 50 55 60

Arg Ile Val Glu Ala Thr Asn Gly His Ser Leu Val Glu Cys Glu Gly 65 70 75 80

Asp Val Val Ile Pro Cys Arg Phe Val Thr Val Ala Ser Gly Ala Ala 85 90 95

Ser Gly Lys Phe Leu Gln Tyr Glu Leu Gly Gly Pro Arg Val Ser Val 100 105 110

Gln Thr Ala Tyr Gly Val Glu Val Glu Val Asp Asn Asn Pro Phe Asp 115 120 125

Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp Tyr Val Arg His Asp 130 135 140

Ala Gln Ser Leu Glu Ala Lys Tyr Pro Thr Phe Leu Tyr Ala Met Pro 145 150 155 160

Met Ser Pro Thr Arg Val Phe Phe Glu Glu Thr Cys Leu Ala Ser Lys 165 170 175

Asp Ala Met Pro Phe Asp Leu Leu Lys Lys Lys Leu Met Leu Arg Leu 180 185 190

Asn Thr Leu Gly Val Arg Ile Lys Glu Ile Tyr Glu Glu Glu Trp Ser 195 200 205

Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Thr Leu 210 215 220

Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser 225 230 235 240

Val Val Arg Ser Leu Ser Glu Ala Pro Lys Cys Ala Phe Val Leu Ala 245 250 255

Asn Ile Leu Arg Gln Asn His Ser Lys Asn Met Leu Thr Ser Ser Ser 260 265 270

Thr Pro Ser Ile Ser Thr Gln Ala Trp Asn Thr Leu Trp Pro Gln Glu 275 280 285

Arg Lys Arg Gln Arg Ser Phe Phe Leu Phe Gly Leu Ala Leu Ile Leu 290 295 300

Gln Leu Asp Ile Glu Gly Ile Arg Ser Phe Phe Arg Ala Phe Phe Arg 305 310 315 320

Val Pro Lys Trp Met Trp Gln Gly Phe Leu Gly Ser Ser Leu Ser Xaa 325 330 335

Ala Asp Leu Met Leu Phe Ala Phe Tyr Met Phe Ile Ile Ala Pro Asn 340 345 350

Asp Met Arg Arg Gly Leu Ile Arg His Leu Leu Ser Asp Pro Thr Gly 355 360 365

Ala Thr Leu Ile Arg Thr Tyr Leu Thr Phe 370 375

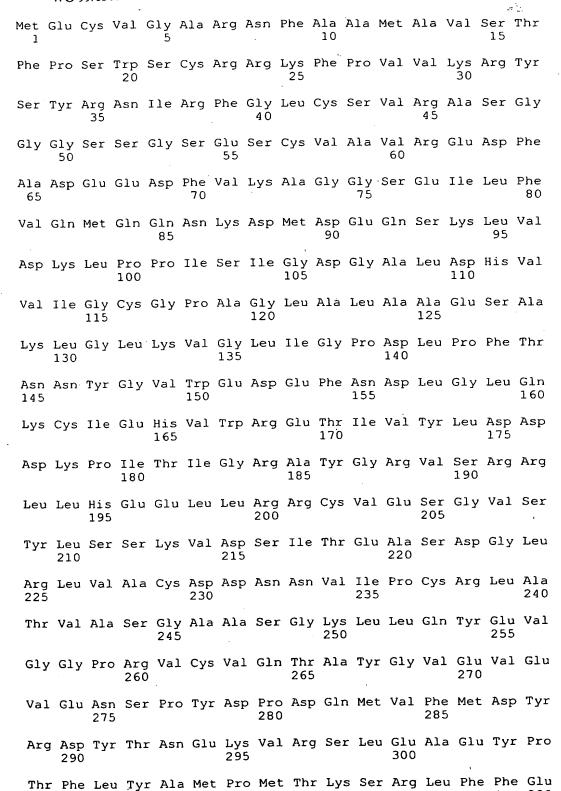
<210> 49

<211> 524

<212> PRT

<213> Arabidopsis thaliana

<400> 49



Glu Thr Cys Leu Ala Ser Lys Asp Val Met Pro Phe Asp Leu Leu Lys

325 330 335

Thr Lys Leu Met Leu Arg Leu Asp Thr Leu Gly Ile Arg Ile Leu Lys 340 345 350

Thr Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro 355 360 365

Asn Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val $370 \hspace{1.5cm} 375 \hspace{1.5cm} 380$

His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro 385 390 395 400

Lys Tyr Ala Ser Val Ile Ala Glu Ile Leu Arg Glu Glu Thr Thr Lys
405
410
415

Gln Ile Asn Ser Asn Ile Ser Arg Gln Ala Trp Asp Thr Leu Trp Pro 420 425 430

Pro Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu Phe Gly Leu Ala Leu 435 440 445

Ile Val Gln Phe Asp Thr Glu Gly Ile Arg Ser Phe Phe Arg Thr Phe 450 455 460

Phe Arg Leu Pro Lys Trp Met Trp Gln Gly Phe Leu Gly Ser Thr Leu 465 470 475 480

Thr Ser Gly Asp Leu Val Leu Phe Ala Leu Tyr Met Phe Val Ile Ser 485 490 495

Pro Asn Asn Leu Arg Lys Gly Leu Ile Asn His Leu Ile Ser Asp Pro 500 505 510

Thr Gly Ala Thr Met Ile Lys Thr Tyr Leu Lys Val 515 520

<210> 50

<211> 529

<212> PRT

<213> Adonis palaestina

<400> 50

Met Glu Leu Cly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp

1 5 . 10 15

Thr Phe Gly Thr Arg Asn Leu Ser Ser Ser Lys Leu Ala Tyr Asn Ile 20 25 30

His Arg Tyr Gly Ser Ser Cys Arg Val Asp Phe Gln Val Arg Ala Asp 35 40 45

Gly Gly Ser Gly Ser Arg Ser Ser Val Ala Tyr Lys Glu Gly Phe Val
50 55 60

Asp Glu Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu Leu Leu Phe Val 65 70 75 80

Gln Met Gln Gln Thr Lys Ser Met Glu Lys Gln Ala Lys Leu Ala Asp

BNSDOCID- -WO GOSSOESA LIA-

85

90 95

Lys Leu Pro Pro Ile Pro Phe Gly Glu Ser Val Met Asp Leu Val Val 105 Ile Gly Cys Gly Pro Ala Gly Leu Ser Leu Ala Ala Glu Ala Ala Lys Leu Gly Leu Lys Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Glu Arg 150 Cys Ile Glu His Ala Trp Lys Asp Thr Ile Val Tyr Leu Asp Asn Asp 170 Ala Pro Val Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu His Glu Glu Leu Leu Lys Arg Cys Val Glu Ser Gly Val Ser Tyr Leu Asp Ser Lys Val Glu Arg Ile Thr Glu Ala Gly Asp Gly His Ser Leu Val Val Cys Glu Asn Glu Ile Phe Ile Pro Cys Arg Leu Ala Thr Val Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Glu Tyr Glu Val Gly Gly Pro Arg Val Cys Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val 265 Glu Asn Asn Pro Tyr Asp Pro Asn Leu Met Val Phe Met Asp Tyr Arg 280 Asp Tyr Met Gln Gln Lys Leu Gln Cys Ser Glu Glu Glu Tyr Pro Thr Phe Leu Tyr Val Met Pro Met Ser Pro Thr Arg Leu Phe Phe Glu Glu 310 315 Thr Cys Leu Ala Ser Lys Asp Ala Met Pro Phe Asp Leu Leu Lys Arg Lys Leu Met Ser Arg Leu Lys Thr Leu Gly Ile Gln Val Thr Lys Val 345 Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn 360 Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys 390 395 Tyr Ala Ser Val Ile Ala Lys Ile Leu Lys Gln Asp Asn Ser Ala Tyr

Val Val Ser Gly Gln Ser Ser Ala Val Asn Ile Ser Met Gln Ala Trp 420 425 430

Ser Ser Leu Trp Pro Lys Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu 435 440 445

Phe Gly Leu Glu Leu Ile Val Gln Leu Asp Ile Glu Ala Thr Arg Thr 450 455 460

Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp Trp Gly Phe 465 470 480

Leu Gly Ser Ser Leu Ser Ser Phe Asp Leu Val Leu Phe Ser Met Tyr 485 490 495

Met Phe Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His 500 505 510

Leu Leu Ser Asp Pro Ser Gly Ala Val Met Val Arg Ala Tyr Leu Glu 515 520 525

Arg

<210> 51

<211> 529

<212> PRT

<213> Adonis palaestina

<400> 51

RNSDOCID: WO GORGOSSA1 IAS

Met Glu Leu Leu Gly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp
1 5 10 15

Thr Phe Gly Thr Arg Asn Leu Ser Ser Ser Lys Leu Ala Tyr Asn Ile 20 25 30

His Arg Tyr Gly Ser Ser Cys Arg Val Asp Phe Gln Val Arg Ala Asp 35 40 45

Gly Gly Ser Gly Ser Arg Thr Ser Val Ala Tyr Lys Glu Gly Phe Val
50 60

Asp Glu Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu Leu Leu Phe Val 65 70 75 80

Gln Met Gln Gln Thr Lys Ser Met Glu Lys Gln Ala Lys Leu Ala Asp 85 90 95

Lys Leu Pro Pro Ile Pro Phe Gly Glu Ser Val Met Asp Leu Val Val 100 105 110

Ile Gly Cys Gly Pro Ala Gly Leu Ser Leu Ala Ala Glu Ala Ala Lys 115 120 125

Leu Gly Leu Lys Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn 130 135 140

Asn Tyr Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Glu Arg 145 150 155 160





| | Cys | Ile | Glu | His | Ala 165 | Trp | Lys | Asp | Thr | Ile 170 | Val | Tyr | Leu | Asp | Asn 175 | |
|-----|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| . 7 | Ala | Pro | Val | Leu 180 | Ile | Gly | Arg | Ala | Tyr 185 | Gly | Arg | Val | Ser | Arg 190 | His | Leu |
|] | Leu | His | Glu 195 | Glu | Leu | Leu | Lys | Arg 200 | Cys | Val | Glu | Ser | Gly 205 | Val | Ser | Tyr |
|] | Leu | Asn 210 | Ser | Lys | Val | Glu | Arg 215 | Ile | Thr | Glu | Ala | Gly 220 | Asp | Gly | His | Ser |
| 1 | Leu 225 | Val | Val | Cys | Glu | Asn 230 | Asp | Ile | Phe | Ile | Pro 235 | Cys | Arg | Leu | Ala | Thr 240 |
| 7 | /al | Ala | Ser | Gly | Ala 245 | Ala | Ser | Gly | Lys | Leu 250 | Leu | Glu | Tyr | Glu | Val 255 | Gly |
| (| Sly | Pro | Arg | Val 260 | Cys | Val | Gln | Thr | Ala 265 | Tyr | Gly | Val | Glu | Val 270 | Glu | Val |
| (| Glu | Asn | Asn 275 | Pro | Tyr | Asp | Pro | Asn 280 | Leu | Met | Val | Phe | Met 285 | Asp | Tyr | Arg |
| 7 | qz | Tyr 290 | Met | Gln | Gln | Lys | Leu 295 | Gln | Cys | Ser | Glu | Glu 300 | Glu | Tyr | Pro | Thr |
| 3 | Phe 305 | Leu | Tyr | Val | Met | Pro 310 | Met | Ser | Pro | Thr | Arg 315 | Leu | Phe | Phe | Glu | Glu 320 |
| 7 | Thr | Cys | Leu | Ala | Ser 325 | Lys | Asp | Ala | Met | Pro 330 | Phe | Asp | Leu | Leu | Lys 335 | Arg |
| Ι | -ys | Leu | Met | Ser 340 | Arg | Leu | Lys | Thr | Leu 345 | Gly | Ile | Gln | Val | Thr 350 | Lys | Ile |
| 7 | ſyr | Glu | Glu 355 | Glu | Trp | Ser | Tyr | Ile 360 | Pro | Val | Gly | Gly | Ser 365 | Leu | Pro | Asn |
| 7 | ľhr | Glu 370 | Gln | Lys | Asn | Leu | Ala 375 | Phe | Gly | Ala | Ala | Ala 380 | Ser | Met | Val | His |
| 3 | Pro 385 | Ala | Thr | Gly | Tyr | Ser 390 | Val | Val | Arg | Ser | Leu 395 | Ser | Glu | Ala | Pro | Lys 400 |
| 7 | ſyr | Ala | Ser | Val | Ile 405 | Ala | Lys | Ile | Leu | Lys 410 | Gln | Asp | Asn | Ser | Ala 415 | Tyr |
| 7 | /al | Val | Ser | Gly 420 | Gln | Ser | Ser | Ala | Val 425 | Asn | Ile | Ser | Met | Gln 430 | Ala | Trp |
| 5 | Ser | Ser | Leu 435 | Trp | Pro | Lys | Glu | Arg 440 | Lys | Arg | Gln | Arg | Ala 445 | Phe | Phe | Leu |
| I | Phe | Gly 450 | Leu | Glu | Leu | Ile | Val 455 | Gln | Leu | Asp | Ile | Glu 460 | Ala | Thr | Arg | Thr |
| 1 | Phe 165 | Phe | Arg | Thr | Phe | Phe 470 | Arg | Leu | Pro | Thr | Trp 475 | Met | Trp | Trp | Gly | Phe 480 |
|] | Leu | Gly | Ser | Ser | Leu | Ser | Ser | Phe | Asp | Leu | Val | Leu | Phe | Ser | Met | Tyr |

485

490

495

Met Phe Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His 500 505 510

Leu Leu Ser Asp Pro Ser Gly Ala Val Met Val Lys Ala Tyr Leu Glu 515 520 525

Arg

<210> 52

<211> 533

<212> PRT

<213> Lettuce

<400> 52

Met Glu Cys Phe Gly Ala Arg Asn Met Thr Ala Thr Met Ala Val Phe 1 5 10 15

Thr Cys Pro Arg Phe Thr Asp Cys Asn Ile Arg His Lys Phe Ser Leu 20 25 30

Leu Lys Gln Arg Arg Phe Thr Asn Leu Ser Ala Ser Ser Ser Leu Arg 35 40 45

Gln Ile Lys Cys Ser Ala Lys Ser Asp Arg Cys Val Val Asp Lys Gln 50 55 60

Gly Ile Ser Val Ala Asp Glu Glu Asp Tyr Val Lys Ala Gly Gly Ser 65 70 75 80

Glu Leu Phe Phe Val Gln Met Gln Arg Thr Lys Ser Met Glu Ser Gln 85 90 95

Ser Lys Leu Ser Glu Lys Leu Ala Gln Ile Pro Ile Gly Asn Cys Ile 100 105 110

Leu Asp Leu Val Val Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala 115 120 125

Ala Glu Ser Ala Lys Leu Gly Leu Asn Val Gly Leu Ile Gly Pro Asp 130 135 140

Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Gln Asp Glu Phe Ile Gly 145 150 155 160

Leu Gly Leu Glu Gly Cys Ile Glu His Ser Trp Lys Asp Thr Leu Val 165 170 175

Tyr Leu Asp Asp Ala Asp Pro Ile Arg Ile Gly Arg Ala Tyr Gly Arg 180 185 190

Val His Arg Asp Leu Leu His Glu Glu Leu Leu Arg Arg Cys Val Glu 195 200 205

Ser Gly Val Ser Tyr Leu Ser Ser Lys Val Glu Arg Ile Thr Glu Ala 210 215 220

Pro Asn Gly Tyr Ser Leu Ile Glu Cys Glu Gly Asn Ile Thr Ile Pro

| | W |) 99/6 | 3055 | Ų | 7 | | | | | | | | | | |
|----------|--------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 22 | 25 | | | | 230 | | | | | 235 | | | | | 240 |
| C7 | 's Ar | j Leu | Ala | Thr 245 | Val | Ala | Ser | Gly | Ala 250 | Ala | Ser | Gly | Lys | Phe 255 | Leu |
| G] | и Туі | Glu | Leu 260 | Gly | Gly | Pro | Arg | Val 265 | Cys | Val | Gln | Thr | Ala 270 | Tyr | Gly |
| 13 | e Glu | Val 275 | Glu | Val | Glu | Asn | Asn 280 | Pro | Tyr | Asp | Pro | Asp 285 | Leu | Met | Val |
| Ph | e Met 290 | Asp | Tyr | Arg | Asp | Phe 295 | Ser | Lys | His | Lys | Pro 300 | Glu | Ser | Leu | Glu. |
| A1 30 | a Lys | Tyr | Pro | Thr | Phe 310 | Leu | Tyr | Val | Met | Ala 315 | Met | Ser | Pro | Thr | Lys 320 |
| 11 | e Phe | Phe | Glu | Glu 325 | Thr | Cys | Leu | Ala | Ser 330 | Arg | Glu | Ala | Met | Pro 335 | Phe |
| As | n Leu | Leu | Lys 340 | Ser | Lys | Leu | Met | Ser 345 | Arg | Leu | Lys | Ala | Met 350 | Gly | Ile |
| Ar | g Ile | Thr 355 | Arg | Thr | Tyr | Glu | Glu 360 | Glu | Trp | Ser | Туr | Ile 365 | Pro | Val | Gly |
| Gl | y Ser 370 | Leu | Pro | Asn | Thr | Glu 375 | Gln | Lys | Asn | Leu | Ala 380 | Phe | Ģly | Ala | Ala |
| A1 38 | a Ser 5 | Met | Val | His | Pro 390 | Ala | Thr | Gly | Tyr | Ser 395 | Val | Val | Arg | Ser | Leu 400 |
| Se | r Glu | Ala | Pro | Asn 405 | Tyr | Ala | Ala | Val | Ile 410 | Ala | Lys | Ile | Leu | Arg 415 | Gln |
| As | p Gln | Ser | Lys 420 | Glu | Met | Ile | Ser | Leu 425 | Gly | Lys | Tyr | Thr | Asn 430 | Ile | Ser |
| Ly | s Gln | Ala 435 | Trp | Glu | Thr | Leu | Trp 440 | Pro | Leu | Glu | Arg | Lys 445 | Arg | Gln | Arg |
| Al | a Phe 450 | Phe | Leu | Phe | Gly | Leu 455 | Ser | His | Ile | Val | Leu 460 | Met | Asp | Leu | Glu |
| G1 46 | y Thr 5 | Arg | Thr | Phe | Phe 470 | Arg | Thr | Phe | Phe | Arg 475 | Leu | Pro | Lys | Trp | Met 480 |
| Tr | p Trp | Gly | Phe | Leu 485 | Gly | Ser | Ser | Leu | Ser 490 | Ser | Thr | Asp | Leu | Ile 495 | Ile |
| Ph | e Ala | Leu | Tyr 500 | Met | Phe | Val | Ile | Ala 505 | Pro | His | Ser | Leu | Arg 510 | Met | Glu |
| Le | u Val | Arg 515 | His | Leu | Leu | Ser | Asp 520 | Pro | Thr | Gly | Ala | Thr 525 | Met | Val | Lys |

<210> 53

Ala Tyr Leu Thr Ile 530

<211> 526 <212> PRT

<213> Tomato

<400> 53 Met Glu Cys Val Gly Val Gln Asn Val Gly Ala Met Ala Val Leu Thr Arg Pro Arg Leu Asn Arg Trp Ser Gly Gly Glu Leu Cys Gln Glu Lys Ser Ile Phe Leu Ala Tyr Glu Gln Tyr Glu Ser Lys Cys Asn Ser Ser Ser Gly Ser Asp Ser Cys Val Val Asp Lys Glu Asp Phe Ala Asp Glu Glu Asp Tyr Ile Lys Ala Gly Gly Ser Gln Leu Val Phe Val Gln Met Gln Gln Lys Lys Asp Met Asp Gln Gln Ser Lys Leu Ser Asp Glu Leu Arg Gln Ile Ser Ala Gly Gln Thr Val Leu Asp Leu Val Val Ile Gly 105 Cys Gly Pro Ala Gly Leu Ala Leu Ala Ala Glu Ser Ala Lys Leu Gly Leu Asn Val Gly Leu Val Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Gln Ala Cys Ile 150 Glu His Val Trp Arg Asp Thr Ile Val Tyr Leu Asp Asp Glu Pro Ile Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Phe Leu His Glu Glu Leu Leu Lys Arg Cys Val Glu Ala Gly Val Leu Tyr Leu Asn Ser Lys Val Asp Arg Ile Val Glu Ala Thr Asn Gly Gln Ser Leu Val Glu Cys Glu Gly Asp Val Val Ile Pro Cys Arg Phe Val Thr Val Ala 235 Ser Gly Ala Ala Ser Gly Lys Phe Leu Gln Tyr Glu Leu Gly Ser Pro 245 Arg Val Ser Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Asp Asn 265 Asn Pro Phe Asp Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp Tyr

280

295

Leu Arg His Asp Ala Gln Ser Leu Glu Ala Lys Tyr Pro Thr Phe Leu



| | | | | | | | | | | | | | | | er 12. |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Tyr 305 | Ala | Met | Pro | Met | Ser 310 | Pro | Thr | Arg | Val | Phe 315 | Phe | Glu | Glu | Thr | Cys 320 |
| Leu | Ala | Ser | Lys | Asp 325 | Ala | Met | Pro | Phe | Asp 330 | Leu | Leu | Lys | Lys | Lys 335 | Leu |
| Met | Leu | Ārg | Leu 340 | Asn | Thr | Leu | Gly | Val 345 | Arg | Ile | Lys | Glu | Ile 350 | Tyr | Glu |
| Glu | Glu | Trp 355 | Ser | Tyr | Ile | Pro | Val 360 | Gly | Gly | Ser | Leu | Pro 365 | Asn | Thr | Glu |
| Gln | Lys 370 | Thr | Leu | Ala | Phe | Gly 375 | Ala | Ala | Ala | Ser | Met 380 | Val | His | Pro | Ala |
| Thr 385 | Gly | Tyr | Ser | Val | Val 390 | Arg | Ser | Leu | Ser | Glu 395 | Ala | Pro | Lys | Cys | Ala 400 |
| Ser | Val | Leu | Ala | Asn 405 | Ile | Leu | Arg | Gln | His 410 | Tyr | Ser | Lys | Asn | Met 415 | Leu |
| Thr | Ser | Ser | Ser 420 | Ile | Pro | Ser | Ile | Ser 425 | Thr | Gln | Ala | Trp | Asn 430 | Thr | Leu |
| Trp | Pro | Gln 435 | Glu | Arg | Lys | Arg | Gln 440 | Arg | Ser | Phe | Phe | Leu 445 | Phe | Gly | Leu |
| Ala | Leu 450 | Ile | Leu | Gln | Leu | Asp 455 | Ile | Glu | Gly | lle | Arg 460 | Ser | Phe | Phe | Arg |
| Ala 465 | Phe | Phe | Arg | Val | Pro 470 | Lys | Trp | Met | Trp | Gln 475 | Gly | Phe | Leu | | Ser 480 |
| Ser | Leu | Ser | Ser | Ala 485 | Asp | Leu | Met | Leu | Phe 490 | Ala | Phe | Tyr | Met | Phe 495 | Ile |
| Ile | Ala | Pro | Asn 500 | Asp | Met | Arg | Lys | Gly 505 | Leu | Ile | Arg | His | Leu 510 | Leu | Ser |
| Asp | Pro | Thr 515 | Gly | Ala | Thr | Leu | Ile 520 | Arg | Thr | Tyr | Leu | Thr 525 | Phe | | |
| | | | | | | | | | | | | | | | |

<210> 54

<211> 516

<212> PRT

<213> Tagetes erecta

<400> 54

Met Ser Met Arg Ala Gly His Met Thr Ala Thr Met Ala Ala Phe Thr $1 \hspace{1cm} 5 \hspace{1cm} 10 \hspace{1cm} 15$

Cys Pro Arg Phe Met Thr Ser Ile Arg Tyr Thr Lys Gln Ile Lys Cys 20 25 30

Asn Ala Ala Lys Ser Gln Leu Val Val Lys Gln Glu Ile Glu Glu Glu 35 40 45

Glu Asp Tyr Val Lys Ala Gly Gly Ser Glu Leu Leu Phe Val Gln Met 50 55 60

BNSDOCID: -WO GORGOERA 1 IA-

| Gln 65 | Gln | Asn | Lys | Ser | Met 70 | Asp | Ala | Gln | Ser | Ser 75 | Leu | Ser | Gln | Lys | Leu 80 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Pro | Arg | Val | Pro | Ile 85 | Gly | Gly | Gly | Gly | Asp 90 | Ser | Asn | Cys | Ile | Leu 95 | Asp |
| Leu | Val | Val | Ile 100 | Gly | Cys | Gly | Pro | Ala 105 | Gly | Leu | Ala | Leu | Ala 110 | Gly | Glu |
| Ser | Ala | Lys 115 | Leu | Gly | Leu | Asn | Val 120 | Ala | Leu | Ile | Gly | Pro 125 | Asp | Leu | Pro |
| Phe | Thr 130 | Așn | Asn | Tyr | Gly | Val 135 | Trp | Glu | Asp | Glu | Phe 140 | Ile | Gly | Leu | Gly |
| Leu 145 | Glu | Gly | Cys | Ile | Glu 150 | His | Val | Trp | Arg | Asp 155 | Thr | Val | Val | Tyr | Leu 160 |
| Asp | Asp | Asn | Asp | Pro 165 | Ile | Leu | Ile | Gly | Arg 170 | Ala | Tyr | Gly | Arg | Val 175 | Ser |
| Arg | Asp | Leu | Leu 180 | His | Glu | Glu | Leu | Leu 185 | Thr | Arg | Cys | Met | Glu 190 | Ser | Gly |
| Val | Ser | Tyr 195 | Leu | Ser | Ser | Lys | Val 200 | Glu | Arg | Ile | Thr | Glu 205 | Ala | Pro | Asn |
| Gly | Leu 210 | Ser | Leu | Ile | Glu | Cys 215 | Glu | Gly | Asn | Ile | Thr 220 | Ile | Pro | Cys | Arg |
| Leu 225 | Ala | Thr | Val | Ala | Ser 230 | Gly | Ala | Ala | Ser | Gly 235 | Lys | Leu | Leu | Gln | Tyr 240 |
| Glu | Leu | Gly | Gly | Pro 245 | Arg | Val | Cys | Val | Gln 250 | Thr | Ala | Tyr | Gly | Ile 255 | Glu |
| Val | Glu | Val | Glu 260 | Ser | Ile | Pro | Tyr | Asp 265 | Pro | Ser | Leu | Met | Val 270 | Phe | Met |
| Asp | Tyr | Arg 275 | Asp | Tyr | Thr | Lys | His 280 | Lys | Ser | Gln | Ser | Leu 285 | Glu | Ala | Gln |
| Tyr | Pro 290 | Thr | Phe | Leu | Tyr | Val 295 | Met | Pro | Met | Ser | Pro 300 | Thr | Lys | Val | Phe |
| Phe 305 | Glu | Glu | Thr | Cys | Leu 310 | Ala | Ser | Lys | Glu | Ala 315 | Met | Pro | Phe | Glu | Leu 320 |
| Leu | Lys | Thr | Lys | Leu 325 | Met | Ser | Arg | Leu | Lys 330 | Thr | Met | Gly | Ile | Arg 335 | Ile |
| Thr | Lys | Thr | Tyr 340 | Glu | Glu | Glu | Trp | Ser 345 | Tyr | Ile | Pro | Val | Gly 350 | Gly | Ser |
| Leu | Pro | Asn 355 | Thr | Glu | Gln | Lys | Asn 360 | Leu | Ala | Phe | Gly | Ala 365 | Ala | Ala | Ser |
| Met | Val 370 | His | Pro | Ala | Thr | Gly 375 | Tyr | Ser | Val | Val | Arg 380 | Ser | Leu | Ser | Glu |
| Ala | Pro | Asn | Tyr | Ala | Ala | Val | Ile | Ala | Lys | Ile | Leu | Gly | Lys | Gly | Asn |

PCT/US99/12121

390 395 Ser Lys Gln Met Leu Asp His Gly Arg Tyr Thr Thr Asn Ile Ser Lys 405 410 Gln Ala Trp Glu Thr Leu Trp Pro Leu Glu Arg Lys Arg Gln Arg Ala 425 Phe Phe Leu Phe Gly Leu Ala Leu Ile Val Gln Met Asp Ile Glu Gly Thr Arg Thr Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp 455 Trp Gly Phe Leu Gly Ser Ser Leu Ser Ser Thr Asp Leu Ile Ile Phe 475 Ala Phe Tyr Met Phe Ile Ile Ala Pro His Ser Leu Arg Met Gly Leu Val Arg His Leu Leu Ser Asp Pro Thr Gly Gly Thr Met Leu Lys Ala 500 505 Tyr Leu Thr Ile 515 <210> 55 <211> 501 <212> PRT <213> Arabidopsis thaliana <400> 55 Met Asp Thr Leu Leu Lys Thr Pro Asn Lys Leu Asp Phe Phe Ile Pro Gln Phe His Gly Phe Glu Arg Leu Cys Ser Asn Asn Pro Tyr His Ser Arg Val Arg Leu Gly Val Lys Lys Arg Ala Ile Lys Ile Val Ser Ser Val Val Ser Gly Ser Ala Ala Leu Leu Asp Leu Val Pro Glu Thr Lys Lys Glu Asn Leu Asp Phe Glu Leu Pro Leu Tyr Asp Thr Ser Lys Ser Gln Val Val Asp Leu Ala Ile Val Gly Gly Pro Ala Gly Leu Ala Val Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile Asp 105 Pro Ser Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Thr Thr Trp Ser

135

Gly Ala Val Val Tyr Val Asp Glu Gly Val Lys Lys Asp Leu Ser Arg

| 145 | | | | | 150 | | | | | 155 | | | | | 160 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Pro | Tyr | Gly | Arg | Val 165 | Asn | Ārg | Lys | Gln | Leu 170 | Lys | Ser | Lys | Met | Leu 175 | Gln |
| Lys | Cys | Ile | Thr 180 | Asn | Gly | Val | Lys | Phe 185 | His | Gln | Ser | Lys | Val 190 | | Asn |
| Val | Val | His 195 | Glu | Glu | Ala | Asn | Ser 200 | Thr | Val | Val | Cys | Ser 205 | | Gly | Val |
| Lys | Ile 210 | Gln | Ala | Ser | Val | Val 215 | Leu | Asp | Ala | Thr | Gly 220 | Phe | Ser | Arg | Cys |
| Leu 225 | Val | Gln | Tyr | Asp | Lys 230 | Pro | Tyr | Asn | Pro | Gly 235 | Tyr | Gln | Val | Ala | Tyr 240 |
| Gly | Ile | Val | Ala | Glu 245 | Val | Asp | Gly | His | Pro 250 | Phe | Asp | Val | Asp | Lys 255 | Met |
| Val | Phe | Met | Asp 260 | Trp | Arg | Asp | Lys | His 265 | Leu | Asp | Ser | Tyr | Pro 270 | Glu | Leu |
| Lys | Glu | Arg 275 | Asn | Ser | Lys | Ile | Pro 280 | Thr | Phe | Leu | Tyr | Ala 285 | Met | Pro | Phe |
| Ser | Ser 290 | Asn | Arg | Ile | Phe | Leu 295 | Glu | Glu | Thr | Ser | Leu 300 | Val | Ala | Arg | Pro |
| Gly 305 | Leu | Arg | Met | Glu | Asp 310 | Ile | Gln | Glu | Arg | Met 315 | Ala | Ala | Arg | Leu | Lys 320 |
| His | Leu | Gly | Ile | Asn 325 | Val | Lys | Arg | Ile | Glu 330 | Glu | Asp | Glu | Arg | Cys 335 | Val |
| Ile | Pro | Met | Gly 340 | Gly | Pro | Leu | Pro | Val 345 | Leu | Pro | Gln | Arg | Val 350 | Val | Gly |
| Ile | Gly | Gly 355 | Thr | Ala | Gly | Met | Val 360 | His | Pro | Ser | Thr | Gly 365 | Tyr | Met | Val |
| Ala | Arg 370 | Thr | Leu | Ala | Ala | Ala 375 | Pro | Ile | Val | Ala | Asn 380 | Ala | Ile | Val | Arg |
| Tyr 385 | Leu | Gly | Ser | Pro | Ser 390 | Ser | Asn | Ser | Leu | Arg 395 | Gly | Asp | Gln | Leu | Ser 400 |
| Ala | Glu | Val | Trp | Arg 405 | Asp | Leu | Trp | Pro | Ile 410 | Glu | Arg | Arg | Arg | Gln 415 | Arg |
| Glu | Phe | Phe | Cys 420 | Phe | Gly | Met | Asp | Ile 425 | Leu | Leu | Lys | Leu | Asp 430 | Leu | Asp |
| Ala | Thr | Arg 435 | Arg | Phe | Phe | Asp | Ala 440 | Phe | Phe | Asp | Leu | Gln 445 | Pro | His | Tyr |
| Trp | His 450 | Gly | Phe | Leu | Ser | Ser 455 | Arg | Leu | Phe | Leu | Pro 460 | Glu | Leu | Leu | Val |
| Phe 465 | Gly | Leu | Ser | Leu | Phe 470 | Ser | His | Ala | Ser | Asn 475 | Thr | Ser | Arg | Leu | Glu 480 |
| | | | | | | | | | | | | | | | |





Ile Met Thr Lys Gly Thr Val Pro Leu Ala Lys Met Ile Asn Asn Leu 485 490 495

Val Gln Asp Arg Asp 500

<210> 56

<211> 502

<212> PRT

<213> Adonis palaestina

<400> 56

Met Asp Thr Leu Leu Arg Thr His Asn Lys Leu Glu Leu Leu Pro Thr 1 5 10 15

Leu His Gly Phe Ala Glu Lys Gln His Leu Val Ser Thr Ser Lys Leu 20 25. 30

Gln Asn Gln Val Phe Arg Ile Ala Ser Arg Asn Ile His Pro Cys Arg 35 40 45

Asn Gly Thr Val Lys Ala Arg Gly Ser Ala Leu Leu Glu Leu Val Pro 50 55 60

Glu Thr Lys Lys Glu Asn Leu Glu Phe Asp Leu Pro Ala Tyr Asp Pro 65 70 . 75 80

Ser Arg Gly Ile Val Val Asp Leu Ala Val Val Gly Gly Pro Ala 85 90 95

Gly Leu Ala Ile Ala Gln Gln Val Ser Glu Ala Gly Leu Leu Val Cys 100 105 110

Ser Ile Asp Pro Ser Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val 115 120 125

Trp Val Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Thr 130 135 140

Thr Trp Ser Gly Ala Val Val Tyr Thr Asp Asp Asn Ser Lys Lys Tyr 145 150 155 160

Leu Asp Arg Pro Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys 165 170 175

Met Leu Gln Lys Cys Val Thr Asn Gly Val Lys Phe His Gln Ala Lys 180 185 190

Val Ile Lys Val Ile His Glu Glu Ser Lys Ser Leu Leu Ile Cys Asn 195 200 205

Asp Gly Ile Thr Ile Asn Ala Thr Val Val Leu Asp Ala Thr Gly Phe 210 215 220

Ser Arg Cys Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln 225 230 235 240

Val Ala Tyr Gly Ile Met Ala Glu Val Glu Glu His Pro Phe Asp Leu 245 250 255



- Asp Lys Met Leu Phe Met Asp Trp Arg Asp Ser His Leu Asn Glu Lys 260 265 270
- Leu Glu Leu Lys Asp. Lys Asn Arg Lys Ile Pro Thr Phe Leu Tyr Ala 275 280 285
- Met Pro Phe Ser Ser Thr Lys Ile Phe Leu Glu Glu Thr Ser Leu Val 290 295 300
- Ala Arg Pro Gly Leu Arg Phe Glu Asp Ile Gln Glu Arg Met Val Ala 305 310 315 320
- Arg Leu Lys His Leu Gly Ile Lys Val Lys Ser Ile Glu Glu Asp Glu 325
- Arg Cys Val Ile Pro Met Gly Gly Pro Leu Pro Val Leu Pro Gln Arg 340 345 350
- Val Val Gly Ile Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly 355 360 365
- Tyr Met Val Ala Arg Thr Leu Ala Ala Pro Val Val Ala Lys Ser 370 375 380
- Ile Val Gln Tyr Leu Gly Ser Asp Arg Ser Leu Ser Gly Asn Glu Leu 385 390 395 400
- Ser Ala Glu Val Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Gln 405 410 415
- Arg Glu Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu 420 425 430
- Gln Gly Thr Arg Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro His 435 440 445
- Tyr Trp His Gly Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Leu 450 460
- Phe Phe Gly Leu Ser Leu Phe Ser His Ala Ser Asn Ala Ser Arg Ile 465 470 475 480
- Glu Ile Met Ala Lys Gly Thr Val Pro Leu Val Asn Met Met Asn Asn 485
- Leu Ile Gln Asp Thr Asp 500
- <210> 57
- <211> 498
- <212> PRT
- <213> Pepper
- <400> 57
- Met Asp Thr Leu Leu Arg Thr Pro Asn Asn Leu Glu Phe Leu His Gly
 1 5 10 15
- Phe Gly Val Lys Val Ser Ala Phe Ser Ser Val Lys Ser Gln Lys Phe 20 25 30



| Gly | Ala | Lys 35 | Lys | Phe | Cys | Glu | Gly 40 | Leu | Gly | Ser | Arg | Ser 45 | Val | Cys | Val |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Lys | Ala 50 | Ser | Ser | Ser | Ala | Leu 55 | Leu | Glu | Leu | Val | Pro 60 | Glu | Thr | Lys | Lys |
| Glu 65 | Asn | Leu | Asp | Phe | Glu 70 | Leu | Pro | Met | Туr | Asp 75 | Pro | Ser | Lys | Gly · | Val 80 |
| Val | Val | Asp | Leu | Ala 85 | Val | Val | Gly | Gly | Gly 90 | Pro | Ala | Gly | Leu | Ala 95 | Val |
| Ala | Gln | Gln | Val 100 | Ser | Glu | Ala | Gly | Leu 105 | Ser | Val | Cys | Ser | Ile 110 | Asp | Pro |
| Asn | Pro | Lys 115 | Leu | Ile | Trp | Pro | Asn 120 | Asn | Tyr | Gly | Val | Trp 125 | Val | Asp | Glu |
| Phe | Glu 130 | Ala | Met | Asp | Leu | Leu 135 | Asp | Cys | Leu | Asp | Ala 140 | Thr | Trp | Ser | Gly |
| Ala 145 | Ala | Val | Tyr | Ile | Asp 150 | Asp | Lys | Thr | Thr | Lys 155 | Asp | Leu | Asn | Arg | Pro 160 |
| Tyr | Gly | Arg | Val | Asn 165 | Arg | Lys | Gln | Leu | Lys 170 | Ser | Lys | Met | Met | Gln 175 | Lys |
| Cys | Ile | Leu | Asn 180 | Gly | Val | Lys | Phe | His 185 | Gln | Ala | Lys | Val | Ile 190 | Lys | Val |
| Ile | His | Glu 195 | Glu | Ser | Lys | Ser | Met 200 | Leu | Ile | Cys | Asn | Asp 205 | Gly | Ile | Thr |
| Ile | Gln 210 | Ala | Thr | Val | Val | Leu 215 | Asp | Ala | Thr | Gly | Phe 220 | Ser | Arg | Ser | Leu |
| Val 225 | Gln | Tyr | Asp | Lys | Pro 230 | Tyr | Asn | Pro | Gly | Tyr 235 | Gln | Val | Ala | Tyr | Gly 240 |
| Ile | Leu | Ala | Glu | Val 245 | ·Glu | Glu | His | Prọ | Phe 250 | Asp | Val | Asn | Lys | Met 255 | Val |
| Phe | Met | Asp | Trp 260 | Arg | Asp | Ser | His | Leu 265 | Lys | Asn | Asn | Val | Glu 270 | Leu | Lys |
| Glu | Arg | Asn 275 | Ser | Arg | Ile | Pro | Thr 280 | Phe | Leu | Tyr | Ala | Met 285 | Pro | Phe | Ser |
| Ser | Asn 290 | Arg | Ile | Phe | Leu | Glu 295 | Glu | Thr | Ser | Leu | Val 300 | Ala | Arg | Pro | Gly |
| Leu 305 | Gly | Met | Asp | Asp | Ile 310 | Gln | Glu | Arg | Met | Val 315 | | Arg | Ļeu | Ser | His 320 |
| Leu | Gly | Ile | Lys | Val 325 | | Ser | Ile | Glu | Glu 330 | Asp | Glu | His | Cys | Val 335 | Ile |
| Pro | Met | Gly | Gly 340 | | Leu | Pro | Val | Leu 345 | Pro | Gln | Arg | Val | Val 350 | Gly | Ile |
| Gly | Gly | Thr | Ala | Gly | Met | Val | His | Pro | Ser | Thr | Gly | Tyr | Met | Val | Ala |

÷ ::

355 360 365

Arg Thr Leu Ala Ala Ala Pro Val Val Ala Asn Ala Ile Ile Gln Tyr 370 375 380

Leu Ser Ser Glu Arg Ser His Ser Gly Asp Glu Leu Ser Ala Ala Val 385 390 395 400

Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Arg Gln Arg Glu Phe Phe
405 410 415

Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu Pro Ala Thr Arg. 420 425 430

Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp His Gly 435 440 445

Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Ile Val Phe Gly Leu 450 455 460

Ser Leu Phe Ser His Ala Ser Asn Thr Ser Arg Leu Glu Ile Met Thr 465 470 475 480

Lys Gly Thr Leu Pro Leu Val His Met Ile Asn Asn Leu Leu Gln Asp 485 490 495

Lys Glu

<210> 58

<211> 500

<212> PRT

<213> Tomato

<400> 58

Met Asp Thr Leu Leu Lys Thr Pro Asn Asn Leu Glu Phe Leu Asn Pro 1 5 10 15

His His Gly Phe Ala Val Lys Ala Ser Thr Phe Arg Ser Glu Lys His 20 25 30

His Asn Phe Gly Ser Arg Lys Phe Cys Glu Thr Leu Gly Arg Ser Val 35 40 45

Cys Val Lys Gly Ser Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr
50 60

Lys Lys Glu Asn Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys
65 70 75 80

Gly Val Val Val Asp Leu Ala Val Val Gly Gly Gly Pro Ala Gly Leu 85 90 95

Ala Val Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile 100 105 110

Asp Pro Asn Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val 115

Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp



| | 130 | | | • | | 135 | | | | | 140 | | | | ar Es |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Ser 145 | Gly | Ala | Ala | Val | Tyr 150 | Ile | Asp | Asp | Asn | Thr 155 | Ala | Lys | Asp | Leu | His 160 |
| Arg | Pro | Tyr | Gly | Arg 165 | Val | Asn | Arg | Lys | Gln 170 | Leu | Lys | Ser | Lys | Met 175 | Met |
| Gln | Lys | Cys | Ile 180 | Met | Asn | Gly | Val | Lys 185 | Phe | His | Gln | Ala | Lys 190 | Val | Ile |
| Lys | Val | Ile 195 | His | Glu | Glu | Ser | Lys 200 | Ser | Met | Leu | Ile | Cys 205 | Asn | Asp | Gly |
| Ile | Thr 210 | Ile | Gln | Ala | Thr | Val 215 | Val | Leu | Asp | Ala | Thr 220 | Gly | Phe | Ser | Arg |
| Ser 225 | Leu | Val | Gln | Tyr | Asp 230 | Lys | Pro | Tyr | Asn | Pro 235 | Gly | Tyr | Gln | Val | Ala 240 |
| Tyr | Gly | Ile | Leu | Ala 245 | Glu | Val | Glu | Glu | His 250 | Pro | Phe | Asp | Val | Asn 255 | Lys |
| Met | Val | Phe | Met 260 | Asp | Trp | Arg | Asp | Ser 265 | His | Leu | Lys | Asn | Asn 270 | Thr | Asp |
| Leu | Lys | Glu 275 | Arg | Asn | Ser | Arg | Ile 280 | Pro | Thr | Phe | Leu | Tyr 285 | Ala | Met | Pro |
| Phe | Ser 290 | Ser | Asn | Arg | Ile | Phe 295 | Leu | Glu | Glu | Thr | Ser 300 | Leu | Val | Ala | Arg |
| Pro 305 | Gly | Leu | Arg | Ile | Asp 310 | Asp | Ile | Gln | Glu | Arg 315 | Met | Val | Ala | Arg | Leu 320 |
| Asn | His | Leu | Gly | Ile 325 | Lys | Val | Lys | Ser | Ile 330 | Glu | Glu | Asp | Glu | His 335 | Cys |
| Leu | Ile | Pro | Met 340 | Gly | Gly | Pro | Leu | Pro 345 | Val | Leu | Pro | Gln | Arg 350 | Val | Val |
| Gly | Ile | Gly 355 | Gly | Thr | Ala | Gly | Met 360 | Val | His | Pro | Ser | Thr 365 | Gly | Tyr | Met |
| Val | Ala 370 | Arg | Thr | Leu | Ala | Ala 375 | Ala | Pro | Val | Val | Ala 380 | Asn | Ala | Ile | Ile |
| Gln 385 | Tyr | Leu | Gly | Ser | Glu 390 | Arg | Ser | His | Ser | Gly 395 | Asn | Glu | Leu | Ser | Thr 400 |
| Ala | Val | Trp | Lys | Asp 405 | Leu | Trp | Pro | Ile | Glu 410 | Arg | Arg | Arg | Gln | Arg 415 | Glu |
| Phe | Phe | Cys | Phe 420 | Gly | Met | Asp | Ile | Leu 425 | Leu | Lys | Leu | Asp | Leu 430 | Pro | Ala |
| Thr | Arg | Arg 435 | Phe | Phe | Asp | Ala | Phe 440 | Phe | Asp | Leu | Glu | Pro 445 | Arg | Tyr | Trp |
| His | Gly 450 | Phe | Leu | Ser | Ser | Arg 455 | Leu | Phe | Leu | Pro | Glu 460 | Leu | Ile | Val | Phe |

Lį

Gly Leu Ser Leu Phe Ser His Ala Ser Asn Thr Ser Arg Phe Glu Ile

Met Thr Lys Gly Thr Val Pro Leu Val Asn Met Ile Asn Asn Leu Leu 485 490 495

Gln Asp Lys Glu 500

<210> 59

<211> 500

<212> PRT

<213> Tobacco

<400> 59

BNSDOCID: < WO

006305541 145

Met Asp Thr Leu Leu Lys Thr Pro Asn Lys Leu Glu Phe Leu His Pro 1 5 10

Val His Gly Phe Ser Val Lys Ala Ser Ser Phe Asn Ser Val Lys Pro 20 25 30

His Lys Phe Gly Ser Arg Lys Ile Cys Glu Asn Trp Gly Lys Gly Val 35 40 45

Cys Val Lys Ala Lys Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr 50 55 60

Lys Lys Glu Asn Leu Asp Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys 65 70 75 80

Gly Leu Val Val Asp Leu Ala Val Val Gly Gly Pro Ala Gly Leu
85 90 95

Ala Val Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Val Ser Ile 100 105 110

Asp Pro Ser Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val

Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Ala Thr Trp 130 135

Ser Gly Thr Val Val Tyr Ile Asp Asp Asn Thr Thr Lys Asp Leu Asp 145 150 155 160

Arg Pro Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys Met Met 165 170 175

Gln Lys Cys Ile Leu Asn Gly Val Lys Phe His His Ala Lys Val Ile 180 185 190

Lys Val Ile His Glu Glu Ala Lys Ser Met Leu Ile Cys Asn Asp Gly
195 200 205

Val Thr Ile Gln Ala Thr Val Val Leu Asp Ala Thr Gly Phe Ser Arg 210 215 220

Cys Leu Val Gln Tyr Asp Lys Pro Tyr Lys Pro Gly Tyr Gln Val Ala 225 230 235 240



Tyr Gly Ile Leu Ala Glu Val Glu Glu His 250 Pro Phe Asp Thr Ser Lys 255

Met Val Leu Met 260 Asp Trp Arg Asp Ser His Leu Gly Asn Asn Met Glu Leu Lys Glu Arg Asn Arg Lys Val Pro Thr Phe Leu Tyr Ala Met Pro

Phe Ser Ser Asn Lys Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg 290 295 300

Pro Gly Leu Arg Met Asp Asp Ile Gln Glu Arg Met Val Ala Arg Leu 305 310 315 320

Asn His Leu Gly Ile Lys Val Lys Ser Ile Glu Glu Asp Glu His Cys 325 330 335

Val Ile Pro Met Gly Gly Ser Leu Pro Val Ile Pro Gln Arg Val Val 340 345 350

Gly Thr Gly Gly Thr Ala Gly Leu Val His Pro Ser Thr Gly Tyr Met 355 360 365

Val Ala Arg Thr Leu Ala Ala Ala Pro Val Val Ala Asn Ala Ile Ile 370 375 380

His Tyr Leu Gly Ser Glu Lys Asp Leu Leu Gly Asn Glu Leu Ser Ala 385 390 395 400

Ala Val Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Arg Gln Arg Glu 415

Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu Pro Ala 420 425 430

Thr Arg Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp 435 440 445

His Gly Phe Leu Ser Ser Arg Leu Tyr Leu Pro Glu Leu Ile Phe Phe 450 455 460

Gly Leu Ser Leu Phe Ser Arg Ala Ser Asn Thr Ser Arg Ile Glu Ile 465 470 475 480

Met Thr Lys Gly Thr Leu Pro Leu Val Asn Met Ile Asn Asn Leu Leu 485 490 495

Gln Asp Thr Glu . 500

<210> 60

<211> 511

<212> PRT

<213> Tagetes erecta

<400> 60

Met Asp Thr Phe Leu Arg Thr Tyr Asn Ser Phe Glu Phe Val His Pro 1 5 10 15

Brighton Smo

Ser Asn Lys Phe Ala Gly Asn Leu Asn Asn Leu Asn Gln Leu Asn Gln 20 25 30

Ser Lys Ser Gln Phe Gln Asp Phe Arg Phe Gly Pro Lys Lys Ser Gln 35 40 45

Phe Lys Leu Gly Gln Lys Tyr Cys Val Lys Ala Ser Ser Ser Ala Leu 50 60

Leu Glu Leu Val Pro Glu Ile Lys Lys Glu Asn Leu Asp Phe Asp Leu 65 70 75 80

Pro Met Tyr Asp Pro Ser Arg Asn Val Val Val Asp Leu Val Val Val Val 95

Gly Gly Gly Pro Ser Gly Leu Ala Val Ala Gln Gln Val Ser Glu Ala 100 105 110

Gly Leu Thr Val Cys Ser Ile Asp Pro Ser Pro Lys Leu Ile Trp Pro 115 120 125

Asn Asn Tyr Gly Val Trp Val Asp Glu Phe Glu Ala Met Asp Leu Leu 130 135 140

Asp Cys Leu Asp Thr Thr Trp Ser Ser Ala Val Val Tyr Ile Asp Glu 145 150 155 160

Lys Ser Thr Lys Ser Leu Asn Arg Pro Tyr Ala Arg Val Asn Arg Lys 165 170 175

Gln Leu Lys Thr Lys Met Leu Gln Lys Cys Ile Ala Asn Gly Val Lys 180 185 190

Phe His Gln Ala Lys Val Ile Lys Val Ile His Glu Glu Leu Lys Ser 195 200 205

Leu Leu Ile Cys Asn Asp Gly Val Thr Ile Gln Ala Thr Leu Val Leu 210 215 220

Asp Ala Thr Gly Phe Ser Arg Ser Leu Val Gln Tyr Asp Lys Pro Tyr 225 230 235 240

Asn Pro Gly Tyr Gln Val Ala Tyr Gly Ile Leu Ala Glu Val Glu Glu 245 250 255

His Pro Phe Asp Val Asp Lys Met Leu Phe Met Asp Trp Arg Asp Ser 260 265 270

His Leu Asp Gln Asn Leu Glu Ile Lys Ala Arg Asn Ser Arg Ile Pro 275 280 285

Thr Phe Leu Tyr Ala Met Pro Phe Ser Ser Thr Arg Ile Phe Leu Glu 290 295 300

Glu Thr Ser Leu Val Ala Arg Pro Gly Leu Lys Met Glu Asp Ile Gln 305 310 315 320

Glu Arg Met Ala Tyr Arg Leu Lys His Leu Gly Ile Lys Val Lys Ser 325 330 335

Ile Glu Glu Asp Glu Arg Cys Val Ile Pro Met Gly Gly Pro Leu Pro

Val Leu Pro Gln Arg Val Leu Gly Ile Gly Gly Thr Ala Gly Met Val 355

His Pro Ser Thr Gly Tyr Met Val Ala Arg Thr Leu Ala Ala Ala Pro 370

Val Ala Lys Ser Ile Ile Arg Tyr Leu Asn Asn Glu Lys Ser Met 390

Val Ala Asp Val Thr Gly Asp Asp Leu Ala Ala Gly Ile Trp Arg Glu 415

Leu Trp Pro Ile Glu Arg Arg Arg Glo Arg Glu Phe Phe Cys Phe Cly

Leu Trp Pro Ile Glu Arg Arg Gln Arg Glu Phe Phe Cys Phe Gly 420 425 430

Met Asp Ile Leu Leu Lys Leu Asp Leu Glu Gly Thr Arg Arg Phe Phe 435 440 445

Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp His Gly Phe Leu Ser 450 460

Ser Arg Leu Phe Leu Pro Glu Leu Val Thr Phe Gly Leu Ser Leu Phe 465 470 475 480

Gly His Ala Ser Asn Thr Cys Arg Val Glu Ile Met Ala Lys Gly Thr 485 490 495

Leu Pro Leu Ala Thr Met Ile Gly Asn Leu Val Arg Asp Arg Glu 500 505 510

<210> 61

<211> 503

<212> PRT

<213> Daffodil

<400> 61

Met Asp Thr Leu Leu Arg Thr His Asn Arg Leu Glu Leu Leu Tyr Pro 1 5 10 15

Leu His Glu Leu Ala Lys Arg His Phe Leu Ser Pro Ser Pro Asn Pro 20 25 30

Gln Asn Pro Asn Phe Lys Phe Phe Ser Arg Lys Pro Tyr Gln Lys Lys 35 40 45

Cys Arg Asn Gly Tyr Ile Gly Val Ser Ser Asn Gln Leu Leu Asp Leu 50 55 60

Val Pro Glu Ile Lys Lys Glu His Leu Glu Phe Asp Leu Pro Leu Tyr 65 70 75 80

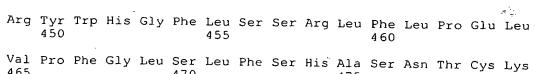
Asp Pro Ser Lys Ala Leu Thr Leu Asp Leu Ala Val Gly Gly Gly 85 90 95

Pro Leu Ala Arg Ser Cys Ser Thr Ser Leu Gly Gly Gly Leu Ser Val

Val Ser Ile Asp Pro Asn Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly

BNSDOCID: < WO

115 120 Val Trp Val Asp Glu Phe Glu Asp Met Asp Leu Leu Asp Cys Leu Asp 135 Ala Thr Trp Ser Gly Ala Ile Val Tyr Val Asp Asp Arg Ser Thr Lys Asn Leu Ser Arg Pro Tyr Ala Arg Val Asn Arg Lys Asn Leu Lys Ser Lys Met Met Lys Lys Cys Val Ser Asn Gly Val Arg Phe His Gln Ala Thr Val Val Lys Ala Met His Glu Glu Glu Lys Ser Tyr Leu Ile Cys 200 Ser Asp Gly Val Thr Ile Asp Ala Arg Val Val Leu Asp Ala Thr Gly Phe Ser Arg Cys Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln Val Ala Tyr Gly Ile Leu Ala Glu Val Glu Glu His Pro Phe Asp 245 Val Asp Lys Met Val Phe Met Asp Trp Arg Asp Ser His Leu Asn Gly 265 Lys Ala Glu Leu Asn Glu Arg Asn Ala Lys Ile Pro Thr Phe Leu Tyr 280 Ala Met Pro Phe Ser Ser Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu 295 Val Ala Arg Pro Gly Leu Lys Met Glu Asp Ile Gln Glu Arg Met Val 315 Ala Arg Leu Asn His Leu Gly Ile Arg Ile Lys Ser Ile Glu Glu Asp Glu Arg Cys Val Ile Pro Met Gly Gly Pro Leu Pro Val Ile Pro Gln 345 Arg Val Val Gly Ile Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met Val Ala Arg Thr Leu Ala Ala Ala Pro Ile Val Ala Asn 375 Ser Ile Val Gln Tyr Leu Val Ser Asp Ser Gly Leu Ser Gly Asn Asp Leu Ser Ala Asp Val Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg 410 Gln Arg Glu Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp 425 Leu Glu Gly Thr Arg Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro 440



Leu Glu Ile Met Ala Lys Gly Thr Leu Pro Leu Val Asn Met Ile Asn 485 490 495

475

Asn Leu Val Gln Asp Arg Asp 500

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/12121

| | FCI70S | 73/12121 |
|--|---|--|
| A. CLASSIFICATION OF SUBJECT MATTER IPC(6): Please See Extra Sheet. US CL: 435/189, 193, 233, 252.3, 320.1, 325; 536/23.2 According to International Patent Classification (IPC) or to both | national classification and IPC | |
| B. FIELDS SEARCHED | national classification and IFC | |
| Minimum documentation searched (classification system follower | d by classification symbols) | |
| U.S. : 435/189, 193, 233, 252.3, 320.1, 325; 536/23.2 | - c,, | |
| Documentation searched other than minimum documentation to th | e extent that such documents are in | cluded in the fields searched |
| Electronic data base consulted during the international search (n Please See Extra Sheet. | ame of data base and, where prac | ticable, search terms used) |
| C. DOCUMENTS CONSIDERED TO BE RELEVANT | | |
| Category* Citation of document, with indication, where ap | propriate, of the relevant passages | Relevant to claim No. |
| WO 97/36998 A1 (UNIVERSITY OF PARK) 09 October 1997, see entire of No:1. | F MARYLAND COLLE | EGE 1-8 9 ID |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| Further documents are listed in the continuation of Box C | See patent family ann | ex. |
| Special categories of cited documents: A document defining the general state of the art which is not considered to be of particular relevance | *T* later document published after date and not in conflict with the principle or theory underly | the international filing date or priority the application but cited to understand ying the invention |
| *E* carlier document published on or after the international filing date | "X" document of particular releva | nce; the claimed invention cannot be |
| *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) | when the document is taken a | considered to involve an inventive step lone nee; the claimed invention cannot be |
| *O* document referring to an oral disclosure, use, exhibition or other means | considered to involve an in- | ventive step when the document is the such documents, such combination |
| *P* document published prior to the international filing date but later than the priority date claimed | *& document member of the same | |
| Date of the actual completion of the international search 02 AUGUST 1999 | Date of mailing of the internation 15 SEP 1999 | al search report |
| Name and mailing address of the ISA.US Commissioner of Patents and Trademarks Box PCT | Authorized officer | as I |
| Washington, D.C. 20231 Facsimile No. (703) 305-3230 | BRADLEÝ S. MAYHEW Telephone No. (703) 368-0196 | 6 Huc |

BNSDOSID WO COSSESSED 210 (second sheet)(July 1992)*

International application No. PCT/US99/12121

| A. CLASSIFICATION | OF | SUBJECT | MATTER: |
|-------------------|----|---------|---------|
| IPC (6): | | | |

C12N 1/21, 5/10, 9/10, 15/53, 15/54, 15/61, 15/63; C12P 23/00; C12Q 1/68

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

Dialog and APS

search terms: 1PP, epsilon cyclase, lycopene cyclase, isopentenyl pyrophosphate isomerase and isopentenyl diphosphate isomerase

THIS PAGE BLANK (USPTO)

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

| Defects in the images include but are not limited to the items chec | ked: |
|---|------|
| ☐ BLACK BORDERS | |
| ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES | |
| ☐ FADED TEXT OR DRAWING | |
| BLURRED OR ILLEGIBLE TEXT OR DRAWING | |
| ☐ SKEWED/SLANTED IMAGES | |
| ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS | |
| ☐ GRAY SCALE DOCUMENTS | |
| ☐ LINES OR MARKS ON ORIGINAL DOCUMENT | \ |
| ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY | |
| OTHER. | |

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)